(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 July 2001 (12.07.2001)

PCT

(10) International Publication Number
WO 01/49832 A2

(51) International Patent Classification[7]: C12N 9/00

(21) International Application Number: PCT/EP01/00060

(22) International Filing Date: 5 January 2001 (05.01.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
00100351.6 7 January 2000 (07.01.2000) EP
00124595.0 10 November 2000 (10.11.2000) EP

(71) Applicant: ARTEMIS PHARMACEUTICALS GMBH [DE/DE]; Neurather Ring 1, 51063 Köln (DE).

(72) Inventor: SCHWENK, Frieder; Kuseler Strasse 4, 50739 Köln (DE).

(74) Agents: HELBING, Jörg et al.; Von Kreisler Selting Werner, Postfach 10 22 41, 50462 Köln (DE).

(81) Designated States *(national)*: AE, AG, AL, AM, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CN, CR, CU, CZ, DM, DZ, EE, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, RO, RU, SD, SG, SI, SK, SL, TJ, TM, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/49832 A2

(54) Title: TRANSDUCTION OF RECOMBINASES FOR INDUCIBLE GENE TARGETING

(57) Abstract: The present invention provides the use of a fusion protein comprising a site-specific DNA recombinase domain and a protein transduction domain for preparing an agent for inducing target gene alteration in a living organism or in cultured cells, suitable fusion proteins and a method for theproduction of said fusion proteins.

Transduction of recombinases for inducible gene targeting

The present invention provides the use of a fusion protein comprising a site-specific DNA recombinase domain and a protein transduction domain for preparing an agent for inducing target gene alteration in a living organism or in cultured cells, suitable fusion proteins and a method for the production of said fusion proteins.

Background

For some years targeted mutagenesis in totipotent mouse embryonic stem (ES) cells has been used to inactivate genes, for which cloned sequences were available (Capecchi, Trends in Genetics 5, 70 - 76 (1989)). Since ES cells can pass mutations induced in vitro to transgenic offspring *in vivo*, it is possible to analyze the consequences of gene disruption in the context of the entire organism. Thus, numerous mouse strains with functionally inactivated genes ("knock out mice") have been created by this technology and utilized to study the biological function of a variety of genes.

A refined method of targeted mutagenesis, referred to as conditional mutagenesis, employs a site-specific recombination system (e.g. Cre/loxP or Flp/frt – Sauer and Henderson, N. Proc. Natl. Acad. Sci. USA *85*, 5166-5170 (1988); Senecoff et al., J. Mol. Biol., 201, 405 - 421 (1988)) which enables a temporally and/or spatially restricted alteration of target genes (Rajewsky et al., J. Clin. Invest., 98, 600 - 603 (1996)). The creation of conditional mouse mutants requires the generation of two mouse strains, i.e. the recombinase recognition strain and the recombinase expressing strain. The recombinase recognition strain is generated by homologous recombination in ES cells as described above except that the targeted

exon(s) is (are) flanked by two recombinase recognition sequences (hereinafter "RRS"; e.g. loxP or frt). The type of recombination event mediated by the recombinase depends on the disposition of the RRS, with deletions, inversions, translocations and integrations being possible (Torres and Kühn, Oxford University Press, Oxford, New York (1997)). By placing the RRS into introns, an interference with gene expression before recombination can be avoided. The recombinase expressing strain contains a recombinase transgene (e.g. Cre, Flp) whose expression is either restricted to certain cells and tissues or is inducible by external agents. Crossing of the recombinase recognition strain with the recombinase expressing strain recombines the RRS-flanked exons from the doubly transgenic offspring in a prespecified temporally and/or spatially restricted manner. Thus, the method allows the temporal analysis of gene function in particular cells and tissues of otherwise widely expressed genes. Moreover, it enables the analysis of gene function in the adult organism by circumventing embryonic lethality which is frequently the consequence of gene mutation. For pharmaceutical research, aiming to validate the utility of genes and their products as targets for drug development, inducible mutations provide an excellent genetic tool. However, the current systems for inducible recombinase expression in transgenic animals suffer from a certain degree of leakiness in the absence of the inducer (Kühn et al., Science 269(5229):1427-9 (1995); Schwenk et al., Nucleic Acids Res.; 26(6):1427-32 (1998)). Furthermore, the generation of conditional mutants is a time consuming and labor intensive procedure, since the recombinase recognition strain and the recombinase expressing strain have to be breed at least over two generations in order to obtain animals carrying both, the recombinase transgene and two copies of the RRS-flanked target gene sequence.

Protein tranduction domains (hereinafter shortly referred to as "PTD") that have the ability to cross cell membranes were identified, e.g. in the

Antennapedia protein from *Drosophila* (Vives et al., J. Biol. Chem, 272(25):16010-7 (1997)), Kaposi fibroblast growth factor (Kaposi FGF; Lin et al., J. Biol. Chem. 270: 14255-58 (1995)), VP22 from HSV (Elliott and O'Hare, Cell, 88(2):223-33 (1997)) and TAT from HIV (Green and Loewenstein, Cell, 55(6):1179-88 (1988); Frankel and Pabo, Cell, 55(6):1189-93 (1988)). WO 99/29721 moreover mentions TAT mutants having an enhanced activity as compared to the wild-type peptide.
Fusion of PTDs to heterologuous proteins conferred the ability to transduce into cultured cells (Fawell et al., Proc. Natl. Acad. Sci. USA, 91(2):664-8 (1994); Elliott and O'Hare (1997), Phelan et al., Nature Biotech. 16; 440-443 (1998) and Dilber et al., Gene Ther., 6(1):12-21 (1999)). Dalby and Bennett showed that a fusion protein consisting of VP22 and functional Flp recombinase translocated between cells in culture (from COS-1 cells transfected with VP22-Flp to CHO cells carrying Flp recognition sites (FRT sites); see Dalby and Bennett, Invitrogen, Expressions 6.2, page 13 (1999)). Further WO 99/11809 mentions a fusion protein Antp-Cre and emphasizes that it may be used to deliver the Cre into the cell which recombines inside the cell nucleus. It is mentioned that the fusion protein is suitable for manipulating genomic DNA at precise locations in a temporal regulated manner.

Furthermore, a recent report demonstrated that the ß-galactosidase protein fused to the 11 amino acids PTD from the HIV TAT protein can infiltrate all tissues of living mice reaching every single cell (Schwarze et al., Science, 285(5433):1569-72 (1999)). Finally, WO 99/60142 discloses vector constructs for gene therapy carrying a tumor cell sensitizing gene, a sensitizing gene expression regulatory system, a control gene and a control gene expression regulatory system, wherein the control gene can be a fusion gene consisting of a recombinase (viz. Cre or Flp) and a trafficking protein (viz. VP22).

With regard to the fusion protein Antp-Cre of WO 99/11809, it is however, general knowledge in the art that the Antennapedia PTD is not a generally applicable transducing protein, namely it has only a limited activity with proteins having more than 100 amino acid residues (Derossi et al., Trends Cell Biol. 8: 84-87, 1998). In view of the limited transducing activity of the Antp PTD and the size of the generally known recombinases (ranging from about 200 to about 600 amino acid residues), it was desirable to provide a more potent system for the transduction of recombinases. It was, however, not clear for a person skilled in the art whether PTDs would be effective at all with recombinases for the following reasons:
(i) only a single example of PTD-mediated delivery of proteins (above 100 amino acid residues) *in vivo* has been reported so far (Schwarze et al., Science, 285(5433):1569-72 (1999); Fawell et al., PNAS, 91: 664-68 (1994); both references describing the TAT-mediated transduction of ß-galactosidase in mice);
(ii) it is known that - due to defolding and refolding processes - the transduction of native proteins into cells may result in a significant loss of protein activity (e.g., as described for TAT-GFP; Schwarze et al, Trends Cell Biol. 10: 290-95 (2000));
(iii) neither the number of protein molecules that can be transferred into a cell by a given translocation domain has been systematically determined, nor the number of Cre molecules in the cell nucleus that is required for efficient recombination;
(iv) the delivery of active proteins requires unfolding- and proper refolding which is unpredictable for a given protein (Bonifaci et al., AIDS 9: 995-1000 1995); and
(v) the mechanism by which protein transduction domains facilitate protein transduction in unknown and several findings have been published that rule out classical receptor-, transporter-, endosome- or endocytosis-mediated processes in the transduction of Ant, TAT and VP22 (G. Eliott, P. O'Hare, Cell 88, 223-233 (1997); D.A. Mann, A.D. Frankel, EMBO. J. 10,

1733-1739 (1991); D. Derossi et al., J. Biol. Chem. 269, 10444-10450 (1994); D. Derossi et al., J. Biol. Chem. 271, 18188-18193 (1996); E. Vives et al., J. Biol. Chem. 272, 16010-16017 (1997)).

Moreover, there was still the need for a generally applicable method where the genetic manipulation can be performed in both, endogenous genes and transgenes.

Summary of the Invention

It was found that site-specific DNA recombinase proteins can be translocated into cells of a living organism when fused to specific protein transduction domains, namely transduction domains being derived from the VP22 protein of HSV or from the TAT protein of HIV. Thus, whenever a gene mutation is desired, recombination is induced upon the injection of the appropriate site-specific recombinase fused to a transduction domain into such a living organism (provided, however, that said organism carries at least one appropriate RRS integrated in the genome).

The present invention thus provides

(1) the use of a fusion protein comprising

(a) a site-specific DNA recombinase domain and

(b) a protein transduction domain (PTD)

for preparing an agent for inducing target gene alterations in a living organism or cell culture, wherein said living organism carries at least one or more recognition sites for said site-specific DNA recombinase integrated in its genome;

(2) a method for inducing gene alterations in a living organism which comprises administering to said living organism a fusion protein comprising a site-specific DNA recombinase domain and a PTD as defined in (1) above, wherein said living organism carries at least one or more

recognition sites for said site-specific DNA recombinase integrated in its genome;

(3) a fusion protein comprising

(a) a site-specific DNA recombinase domain and

(b) a PTD being derived from the VP22 protein of HSV or from the TAT protein of HIV

provided that when the site-specific DNA recombinase domain is wild-type Cre or Flp then the PTD is not the full length VP22 PTD of HSV (i.e., the fusion protein is not identical to the fusion protein of Dalby and Bennett, Invitrogen, Expressions 6.2, page 13 (1999) and of WO 99/60142);

(4) a DNA sequence coding for the fusion protein of (3) above;

(5) a vector comprising the DNA sequence as defined in (4) above;

(6) a host cell transformed with the vector of (5) above and/or comprising the DNA of (4) above;

(7) a method for producing the fusion protein of (1) above which comprises culturing the transformed host cell of (6) above and isolating the fusion protein; and

(8) an injectable composition comprising the fusion protein as defined in (1) or (3) above.

The invention is further illustrated by the appended Figures and is explained in detail below.

Description of the Figures

<u>Fig. 1:</u> Generation of induced mouse mutants using purified fusion proteins.

A: Expression of the fusion protein consisting of the site-specific DNA recombinase (e.g. Cre) and the protein transduction domain (e.g. the HIV derived TAT peptide) in prokaryotic or eukaryotic cells.

B: Extraction and purification of the expressed fusion protein (e.g. as described in Nagahara et al., Nat. Med. 4 (12):1449-52 (1998)).

C: Injection of the purified fusion protein into mice carrying the RRS-flanked target sequence.

D: Analysis of the pattern of induced target gene recombination and the resulting phenotype.

Triangle: RRS.

Fig. 2: Scheme of the bacterial expression vector pT7-TACS (SEQ ID NO:16). The coding region of the 11 amino acid protein transduction domain of HIV TAT protein is fused to the N-terminus of the Cre recombinase protein sequence. The 10-amino-acid strep tag and the protease factor Xa recognition sequence are fused to the C-terminus. The T7 promoter permits expression of TAT-Cre protein in *E. coli*.

Fig. 3: Detection of purified TAT-Cre protein by Coomassie staining and Western blot analysis.

A: Coomassie stained SDS-PAGE gel. Lane 1: 10 kDa ladder (Life Technologies, Cat. No.: 10064-012), 2: 1000 ng BSA, 3: 750 ng BSA, 4: 500 ng BSA, 5: 100 ng BSA, 6: 50 ng BSA, 7: 5 µl TAT-Cre, 8: 1 µl TAT-Cre in Bicine buffer.

B: Western blot analysis using an alkaline phosphatase-conjugated anti-strep tag antibody (IBA, Cat. No: 2-1503-001). Lane 1: MultiMark (Invitrogen, Cat. No.: LC5725), 2: 7 µl TAT-Cre, 3: 5 µl TAT-Cre, 4: 2,5 µl TAT-Cre, 5: 1,25 µl TAT-Cre in Bicine buffer.

Fig. 4: X-Gal staining of M5Pax8 cells treated with TAT-Cre protein. M5Pax8 fibroblasts where treated for 18 h with 3,5 (A), 6,9 (B) and 13,8 µg/ml TAT-Cre protein (C) in serum-free medium. Four days after treatment, cells were fixed and stained with X-Gal.

Fig. 5: Measurement of ß-galactosidase activity in cell lysates. M5Pax8 fibroblasts where treated for 18 h with increasing concentrations of TAT-Cre, as indicated, or transiently transfected with either expression vectors

for Cre (pCMV-I-Cre-pA, see SEQ ID NO:29) or ß-galactosidase (pCMV-I-ß-pA, see SEQ ID NO:30). Four days after treatment, cells were lysed and the ß-galactosidase activities were determined.

<u>Fig. 6:</u> PCR detection of TAT-Cre mediated recombination in mice.
A: PCR-analysis of genomic DNA from duodenum (lane 2), liver (3), kidney (4), spleen (5), muscle (6), lung (7), tail (8) and brain (9) of a *pln13* mouse treated three times with intraperitoneal injections of 75 µg TAT Cre protein at two-day-intervals. Deletion of the loxP-flanked DNA segment is indicated by the presence of the about 400 bp fragment. Lane 1: 1-kb-ladder (Life Technologies).
B: PCR strategy to detect Cre-mediated deletion of the loxP-flanked DNA segment. Arrows indicate the positions of the primers.
C: PCR-analysis of genomic DNA from spleen of a *pln13* mouse treated three times with intraperitoneal injections of 75 µg TAT Cre protein at two-day-intervals (lane 4). To confirm the presence of the BamH I restriction site, the PCR product was digested with BamH I which produces two diagnostic fragments of about 190 and about 210 bp (5). As a control, tail DNA from untreated mice carrying the loxP-flanked (lane 2) and the detected pln13 allele (3) was subjected to PCR amplification. Lane 1: 100 bp ladder (Life Technologies), lane 6: 1 kb ladder (Life Technologies).

<u>Fig. 7:</u> Scheme of the bacterial expression vectors pT7-VPCS (SEQ ID NO:17) and pCRT7-ΔVPCS (SEQ ID NO:15). The coding region of the 301 amino acid protein transduction domain of HSV VP22 protein (A) or the truncated 143 amino acid ΔVP22 domain (B) is fused to the N-terminus of the Cre recombinase protein sequence. The 10-amino-acid strep tag and the protease factor Xa recognition sequence are fused to the C-terminus. The T7 promoter allows the expression of VP22-Cre and ΔVP22-Cre fusion proteins in *E. coli*. The sequence in pCRT7-ΔVPCS encoding the 15 amino

acid N-terminal leader sequence is used for enhanced protein stability (Invitrogen).

Fig. 8: Detection of the purified VP22-Cre and ΔVP22-Cre fusion proteins by Coomassie staining and Western blot analysis.
A: Detection of VP22-Cre protein in a Coomassie-stained SDS-PAGE gel. Lane 1: 10 kDa ladder, 2: 1000 ng BSA, 3: 500 ng BSA, 4: 100 ng BSA, 5: inclusion body protein extract before chromatography, 6: unbound protein, 7: fraction 17, 8: fraction 18, 9: fraction 19, 10: fraction 20. The position of the 75 kDa VP22-Cre protein is indicated by the arrow head.
B: Detection of VP22-Cre protein by Western blot analysis using an alkaline phosphatase-conjugated anti-strep tag antibody (IBA, Cat. No.: 2-1503-001). Lane 1: MultiMark (Invitrogen), 2: inclusion body protein extract before chromatography, 3: unbound protein, 4: fraction 10, 5: fraction 11, 5: fraction 16, 6: fraction 17, 7: fraction 18, 8: fraction 19, 9: fraction 19, 10: fraction 20.
C: Detection of ΔVP22-Cre protein in a Coomassie-stained SDS-PAGE gel. Lane 1: 10 kDa ladder, 2: inclusion body protein extract before chromatography, 3: unbound protein, 4: fraction 1, 5: fraction 8, 6: fraction 9, 7: fraction 15, 8: 100 ng BSA, 9: 500 ng BSA, 10: 1000 ng BSA. The position of the 60 kDa ΔVP22-Cre protein is indicated by the arrow head.
D: Detection of ΔVP22-Cre protein by Western blot analysis using a alkaline phosphatase-conjugated anti-strep tag antibody (IBA, Cat. No.: 2-1503-001). Lane 1: MultiMark (Invitrogen), 2: inclusion body protein extract before chromatography, 3: unbound protein, 4: fraction 4, 5: fraction 8, 6: fraction 10, 7: fraction 12, 8: soluble protein extract before chromatography, 9: unbound protein, 10: fraction 7.

Fig. 9: X-Gal staining of M5Pax8 cells treated with VP22-Cre and ΔVP22-Cre fusion proteins. M5Pax8 fibroblasts where treated for 18 h with either

Bicine buffer (A), 0.5 µg/ml VP22-Cre (B) or 3.75 g/ml ΔVP22-Cre (C) in serum-free medium. Four days after treatment, cells were fixed and stained with X-Gal.

Fig. 10: Measurement of ß-galactosidase activity in cell lysates. M5Pax8 fibroblasts where treated for 18 h with VP22-Cre, ΔVP22-Cre or Bicine buffer alone, as indicated or transiently transfected with expression vectors for Cre (pCMV-I-Cre-pA, see SEQ ID NO:29) or ß-galactosidase (pCMV-I-ß-pA, see SEQ ID NO:30). Four days after treatment, cells were lysed and the ß-galactosidase activities were determined.

Fig. 11: PCR detection of Cre mediated recombination in cells treated with VP22-Cre and ΔVP22-Cre fusion proteins shown in SEQ ID NOs: 21 and 14, respectively).

A: PCR-analysis of genomic DNA isolated from M5Pax8 fibroblasts. Cells were transiently transfected with a Cre expression vector (lane 2) or treated for 18 h with either buffer alone (lane 3), 7.5 µg/ml VP22-Cre (4, 5) or 15 µg/ml ΔVP22-Cre (6, 7) in serum-free medium. Four days after treatment, genomic DNA was extracted and subjected to PCR amplification. Deletion of the loxP-flanked DNA segment is indicated by the presence of the 226 bp DNA fragment. To confirm the presence of the Nco I restriction site in the recombined allele, the PCR products were digested with Nco I which produces two diagnostic fragments of 85bp and 141bp (lanes 5 and 7). Lane 1: 100 bp ladder (Life Technologies), lane 8: 1 kb ladder (Life Technologies).

B: PCR strategy to detect Cre-mediated deletion of the loxP-flanked DNA segment. Arrows indicate the positions of the primers.

Detailed Description of the Invention

The expression "target sequences" according to the present invention means all kind of sequences which may be mutated (viz. deleted,

translocated, integrated and/or inverted) by the action of the recombinase. The number of RRS in the target sequence depends on the kind of mutation to be performed by the recombinase. For most of the mutations (especially for deletions and invertions) two RRS are required which are flanking the sequence to be mutated (deleted or inverted). For some kinds of integrations only one RRS may be necessary within the target sequence.

The "living organisms" according to the present invention are multi-cell organisms and can be vertebrates such as mammals (e.g., rodents such as mice or rats) or non-mammals (e.g., fish) or can be invertebrates such as insects or worms, or can be plants (higher plants, algi or fungi). Most preferred living organisms are mice and fish.

"Cell culture" according to the present invention include cells isolated from the above defined living organism and cultured *in vitro.* These cells can be transformed (immortalized) or untransformed (directly derived from the living organism; primary cell culture).

The site-specific DNA recombinase domain within the fusion protein of the invention of the present application is preferably selected from a recombinase protein derived from Cre, Flp, ϕC31 recombinase (Thorpe and Smith, Proc. Natl. Acad. Sci, USA, vol. 95, 5505-5510 (1998)), $\gamma\delta$ resolvase (Schwickardi and Dröge, FEBS letters 471:147-150 (2000) and R recombinase (Araki et al., J. Mol. Biol., 182, 191-203 (1985)). The preferred recombinases are Cre and mutants thereof (preferably the Cre variant of aa 15 to 357 of SEQ ID NO: 2 or aa 325-667 of SEQ ID NO: 6) and Flp and variants thereof including Flpe (preferably the Flp variant of aa 15 to 437 of SEQ ID NO: 4 or aa 325 to 747 of SEQ ID NO: 8).

The protein transduction domain according to the present invention includes, but is not limited to, the PTDs mentioned in Background of the Invention. The PTD preferably is derived from the VP22 protein of HSV or from the TAT protein of HIV. Suitable TAT proteins include, but are not limited to, proteins comprising (i) the amino acid sequence shown in SEQ ID NO: 10 and mutant thereof such as
(ii) proteins comprising the amino acid
AGRKKRRQRRR (SEQ ID NO:22)
YARKARRQARR (SEQ ID NO:23)
YARAAARQARA (SEQ ID NO:24)
YARAARRAARR (SEQ ID NO:25)
YARAARRAARA (SEQ ID NO:26)
YARRRRRRRRR (SEQ ID NO:27)
YAAARRRRRRR (SEQ ID NO:28)
as known from WO 99/29721. Preferred are transduction domains consisting of the TAT proteins (i) and (ii) above.

Suitable VP22 proteins include, but are not limited to, the wild-type VP22 protein, i.e., a protein comprising amino acids 1 to 302 of SEQ ID No:21, and truncated forms thereof. Truncated VP22 proteins in accordance with the present invention can be those lacking 1 to 158 amino acid residues at their N-terminal end. The most preferred VP22 protein is the truncated VP22 PTD comprising amino acid residues 16 to 157 of SEQ ID NO:14.

The fusion of the two domains of the fusion protein can occur at any possible position, i.e., the protein transduction domain can be fused to the N- or C-terminal of the site-specific DNA recombinase or can be fused to active sites within the site-specific DNA recombinase. Preferably the protein transfusion domain is fused to the N-terminal of the site-specific DNA recombinase domain.

The protein transduction domain can be fused to the site-specific DNA recombinase either through a direct chemical bond or through a linker molecule. Such linker molecule can be any bivalent chemical structure capable of linking the two domains. The preferred linker molecule according to the present invention is a short peptide, e.g., having 1 to 20, preferably 1 to 10, amino acid residues. Specifically preferred short peptides are essentially consisting of Gly, Ala and/or Leu.

The fusion protein of the invention of the present application may further comprise other functional sequences such as secretion conferring signals, nuclear localisation signals and/or signals conferring protein stabilisation.

In case the fusion protein comprises a protein transduction domain derived from the TAT protein of HIV, the DNA sequence coding for said fusion protein preferably comprises the sequence
5' TAC GGC CGC AAG AAG CGC CGC CAA CGC CGC CGC 3'.

Such a preferred DNA sequence is for instance shown in SEQ ID NO: 11. In said sequence the 3' terminal codon ggc codes for the linker Gly. The DNA sequence of a suitable recombinase may be directly attached to said codon ggc.

The fusion protein can be obtained by the following steps:

1. Fusion of the recombinase coding region (e.g. encoding Cre: see amino acids 15 to 357 of SEQ ID NO: 2) with the sequence conferring protein translocation (e.g. the sequence encoding the TAT peptide YGRKKRRQRRR, SEQ ID NO: 10) using standard cloning protocols (Maniatis et al., Cold Spring Harbor Laboratory, New York (1989)) or chemical synthesis.

2. Generation of a construct for the expression of the fusion protein in prokaryotic or eukaryotic cells, e.g. in E. coli DH5a (Hanahan, J. Mol. Biol.;166(4):557-80 (1983)) using the QIAexpress pQE vector (Qiagen, Hilden).
3. Expression of the above mentioned fusion protein in prokaryotic or eukaryotic cells, e.g. in E. coli DH5a (Hanahan, 1983)
4. Extraction and purification of the above mentioned fusion protein e.g. as described in Nagahara et al., Nat. Med., 4(12):1449-52 (1998).

In an experiment it was shown that TAT-mediated delivery of active Cre protein works with sufficient efficacy to facilitate inducible gene targeting both in cell lines and living organisms. In this experiment a vector for the expression of a TAT-Cre fusion protein in E. coli was constructed, TAT-Cre protein was expressed in E. coli and purified from bacterial lysates. To test the activity of the TAT-Cre protein *in vitro*, a reporter cell line that contains a loxP-containing reporter construct was used. This reporter, when recombined by Cre recombinase, allows the expression of a ß-galacosidase gene. Further, a transgenic mouse strain carrying a loxP-flanked target was used to invest the activity of the TAT-Cre protein *in vivo*.

In a second experiment it was shown that VP22-mediated delivery of active Cre protein works with sufficient efficacy to facilitate inducible gene targeting. In this experiment Bacterial expression vectors were constructed for the production of VP22-Cre fusion proteins in E. coli. The activity of purified VP22-Cre proteins were tested using a reporter fibroblast cell line containing a loxP-flanked reporter construct.

Thus, the injection of the purified fusion protein of the present invention into a living organism (e.g., a mouse) carrying a gene comprising the RRS-flanked target sequence (e.g., in an amount of 1 to 200, preferably 5

to 50 µg per g body weight). To demonstrate the feasibility of the invention, a reporter mouse strain carrying an RRS-flanked cassette was used (Thorey et al., Mol. Cell Biol., 18(10):6164 (1998)).

Analysis is achieved by determining the pattern of induced target gene recombination (e.g. through PCR analysis, Southern blot analysis or X-Gal staining on tissue sections; Maniatis et al., 1989; Gossler and Zachgo, Joyner AL (Ed.), Oxford University Press, Oxford, New York (1993)).

The procedure's advantages over current technology are as follows:

(i) The absence of background recombination before administration of the fusion protein.

(ii) The reduction of time and resources which are necessary to combine the recombinase transgene and two copies of the RRS-flanked target gene by conventional breeding.

In experiments it was shown the following: (a) With a suitable vector for the expression of a TAT-Cre fusion protein, a TAT-Cre fusion protein was expressed in *E. coli* and purified from bacterial lysates.
(b) A reporter cell line containing a loxP-containing reporter construct was used to test the activity of the TAT-Cre protein *in vitro*. This reporter, when recombined by Cre recombinase, allows the expression of a ß-galacosidase gene.
(c) A transgenic mouse strain carrying a loxP-flanked target was used to invest the activity of the TAT-Cre protein *in vivo*.
These experiments demonstrate that TAT-mediated delivery of active Cre protein works with sufficient efficacy to facilitate inducible gene targeting both in cell lines and living organisms.

Furthermore, bacterial expression vectors were constructed for the production of VP22-Cre fusion proteins in E. coli. The activity of purified VP22-Cre proteins were tested using a reporter fibroblast cell line containing a loxP-flanked reporter construct. These experiments demonstrate that VP22-mediated delivery of active Cre protein works with sufficient efficacy to facilitate inducible gene targeting.

The invention is further illustrated by the following, non-limitative examples.

Examples

Materials and Methods

Construction of pT7-TACS: The TAT-Cre coding region was generated by PCR using Advantage-HF PCR Kit (Clontech), 20 pmol of the primers TATcre sense (5′-atg cca tgg gct acg gcc gca aga agc gcc gcc aac gcc gcc gcg gca tgt cca att tac tga ccg tac acc-3′; SEQ ID NO:31) and TATcre antisense (5′-ttt cgg atc cgc cgc ata acc agt g-3′; SEQ ID NO:32) and 10 ng pCMV-I-Cre-pA (see SEQ ID NO:29) as template. The PCR reaction was performed using the following cycle profile: 2′ 94 °C, 4 x (30″ 94 °C min, 30″ 50 °C, 1′ 72 °C), 12 x (30″ 94 °C min, 30″ 55 °C, 1′ 72 °C) and 10′ 72 °C. The resulting PCR fragment was digested with Nco I and BamH I, treated with Klenow enzyme and ligated into the plasmid pBSII KS+ which had been opened with restriction enzyme BamH I, treated with Klenow and dephosphorylated with calf intestinal phosphatase. The resulting plasmid pBS TAT-5′cre was verified by DNA sequencing. The Plasmid pCMV-I-Cre-pA (SEQ ID NO:29) was digested with Age I and Sal I which released a 1,036 kb fragment containing the 3′ part of the Cre coding region. This fragment was ligated into the plasmid pBS TAT-5′cre which had been opened with Age I and Sal I.

10 ng pBS-TATCre was subjected to PCR amplification using 20 pmol of primers FPA001 (5'-tat atc tag acc atg ggc tac ggc cgc aag aag c-3'; SEQ ID NO:33) and FPA002 (5'-gct acc acg acc ttc gat acc atc gcc atc ttc cag cag gcg c-3'; SEQ ID NO:34). PCR was performed using 2,5 U Platinum Pfx DNA polymerase (Gibco BRL) and 2 x Enhancer Solution (Gibco BRL) according to the manufacturers protocol. The following cycle profile was used: 2' 94 °C, 25 x (30" 94 °C min, 15" 54,6 °C, 2'30" 68 °C). The amplified PCR fragment was purified using GFX columns (Amersham Pharmacia), digested with Xba I and ligated into the plasmid pASK57 (Skerra and Arne, Gene 151: 131-135 (1994)) which had been opened with restriction enzymes Xba I and Eco 47 III and dephosphorylated with calf intestinal phosphatase. The resulting plasmid pASK75-TACS was digested with restriction enzymes Nco I and Hind III which released a 1,1 kb fragment. The fragment was subsequently ligated into the plasmid pT7-7 (Studier and Moffatt, J. Mol. Biol. 189: 113-130 (1986)) which had been opened with restriction enzymes Nco I and Hind III and dephosphorylated with calf intestinal phosphatase resulting in the plasmid pT7-TACS (SEQ ID NO:16).

Construction of pT7-VPCS: The Cre coding region was generated by PCR using Advantage-HF PCR Kit (Clontech), 20 pmol of the primers VP22cre sense (5'-taa cta gcg gcc gca tgt cca att tac tga ccg tac ac-3'; SEQ ID NO:35) and VP22cre antisense (5'-tcg agc ggc cgc cat cgc cat ctt cca gca ggc g-3'; SEQ ID NO:36) and 10 ng pgkcre-pA (SEQ ID NO:40) as template. The PCR reaction was performed using the following cycle profile: 2' 94 °C, 5 x (30" 94 °C, 30" 50 °C, 2' 72 °C), 15 x (30" 94 °C, 30" 55 °C, 2' 72 °C) and 10' 72 °C. The resulting PCR fragment was digested with Not I and ligated into the plasmid pVP22/Myc-His (Invitrogen), which had been opened with restriction enzyme NotI, dephosphorylated with calf intestinal phosphatase. The resulting plasmid pVP22-cre myc/His was verified by DNA sequencing.

10 ng pVP22-cre myc/His was subjected to PCR amplification using 20 pmol of primers FPA004 (5′-tat atc tag aca tat gac ctc tcg ccg ctc cg-3′; SEQ ID NO:37) and FPA002 (SEQ ID NO:34). PCR was performed using 2,5 U Platinum Pfx DNA polymerase (Gibco BRL) and 2 x Enhancer Solution (Gibco BRL) according to the manufacturers protocol. The following cycle profile was used: 2′ 94 °C, 25 x (30″ 94 °C min, 15″ 54,6 °C, 2′30″ 68 °C). The amplified PCR fragment was purified using GFX columns (Amersham Pharmacia), digested with Xba I and ligated into the plasmid pASK57 (Skerra and Arne, Gene 151: 131-135 (1994)) which had been opened with restriction enzymes Xba I and Eco 47 III and dephosphorylated with calf intestinal phosphatase. The resulting plasmid pASK75-VPCS was digested with restriction enzymes Nde I and Hind III which released a 2,0 kb fragment. The fragment was subsequently ligated into the plasmid pT7-7 (Studier and Moffatt, J. Mol. Biol. 189: 113-130 (1986)) which had been opened with restriction enzymes Nde I and Hind III and dephosphorylated with calf intestinal phosphatase resulting in the plasmid pT7-VPCS (SEQ ID NO:17).

<u>Construction of pCRT7-ΔVPCS:</u> The ΔVP22-Cre coding region was generated by PCR using Platinum Pfx DNA polymerase (Life Technologies), 20 pmol of the primers FPA007 (5'-ttc cga aga cga cga aac acc-3'; SEQ ID NO:38) and FPA008 (5'-tat att cga agc tta tta acc acc gaa ctg cg-3'; SEQ ID NO:39) and 30 ng pT7-VPCS (SEQ ID NO:17) as template. The PCR reaction was performed using the following cycle profile: 2' 94 °C, 25 x (30" 94 °C, 30" 61 °C, 2'30″ 68 °C) and 7' 68 °C. The resulting 1,8 kb PCR fragment was digested with Nco I and Sfu I and ligated into the plasmid pCRT7/VP22-1 (Invitrogen), which had been opened with restriction enzymes Nco I and Sfu I, and dephosphorylated with calf intestinal phosphatase. The resulting plasmid pCRT7-ΔVPCS (SEQ ID NO:15) was verified by DNA sequencing.

<u>Expression of the fusion proteins in E. coli:</u> E. coli BL21(DE3)-RIL cells (Stratagene) were transformed with pT7-TACS and grown on LB agar plates containing 100 µg/ml ampicillin. E. coli BL21(DE3)-RP cells (Stratagene) were transformed with pT7-VPCS and grown on LB agar plates containing 100 µg/ml ampicillin. E. coli BL21(DE3)-pLysS (Invitrogen) were transformed with pCRT7-ΔVPCS and grown on LB agar plates containing 25 µg/ml kanamycine and 34 µg/ml chloramphenicol. Single colonies were isolated and used to prepare glycerol stocks. Eight 5ml LB (Lura Bertani) aliquots containing antibiotics were inoculated with stabs from the glycerol stocks and grown overnight at 37°C with shaking. Two 5ml overnight cultures were each used to inoculate one of four 1L LB aliquots containing antibiotics and grown at 37°C with shaking. Growth rate was monitored by spectrophotometry at 578nm. When the cultures had obtained an OD_{578} = 0,5 expression of the fusion proteins were induced by the addition of 0,5 mM Isopropyl-ß-D-1-thiogalactopyranosid (IPTG). Two hours after induction cells were harvested by centrifugation at 12000xg and the pellet rapidly frozen in liquid nitrogen and stored immediately at −80°C.

<u>Purification of the fusion proteins from bacterial lysates:</u> Each 10g cell pellet was resuspended on ice in 30ml Bicine buffer (50mM Bicine, pH 8,5) including one protease inhibitor tablet (Complete, Roche). Cells were lysed through threefold treatment (1500psi, 5 minutes) with the cell disruption bomb (Parr Instrument). 30ml of Benzonase (10000U, Merck) was added and cell extracts were incubated for 30 minutes at 4°C. Cell extracts were then centrifuged at 12,000xg (4°C). The pellet was redissolved in 8M urea, 50mM Bicine, 100mM DTT, pH 8,5 by incubation for 16 hours at 4°C. Protein extract was centrifuged at 31000xg and supernatant harvested. Protein extract was diluted in an equal volume of Chromatography buffer A (50mM Bicine, pH 8,5). PH was adjusted to pH

8,5 and the extract was filtered through a 0,45µm filter (Millipore). FPLC (Akta Explorer, Amersham Pharmacia) was performed using a cation exchange column (Sepharose SP, Column body HR_5/5 (0.5 x 5cm), column volume (CV) 1ml, linear flow 300cm/hour, Amersham Pharmacia). After addition of sample to FPLC column, buffer was exchanged with Chromatography buffer A at 10 CV.

TAT-Cre and VP22-Cre fusion proteins were eluted from the column by gradient elution using chromatography buffer B (50mM Bicine, 1M NaCl, pH 8,5) using the following profile: 0 - 50 % buffer B, 0 CV; 50 % buffer B, 10 CV; 50 - 100 % buffer B (linear gradient), 20 CV; 100 % buffer B, 10 CV. ΔVP22-Cre protein was eluted from the column by gradient elution using the following profile: 0 - 10 % buffer B, 0 CV; 10 % buffer B, 10 CV; 10 – 30 % buffer B, 0 CV; 30 % buffer B, 10 CV; 30 – 100 % buffer B, 0 CV; 100 % buffer B, 10 CV. Three 1,5ml fractions each containing purified fusion proteins were collected. Purity and concentration of protein fractions were determined by Coomassie blue stained SDS-PAGE gels and Western blot analysis using dilutions of BSA standard solutions. In addition protein content was determined using a Bradford assay (Coomassie Plus protein assay, Pierce).

<u>SDS-PAGE and Western blot analysis:</u> SDS-PAGE and Coomassie staining was performed according to standard protocols (Maniatis et al., Cold Spring Harbor Laboratory, New York (1989)) using 4 - 12 % gradient SDS-polyacrylamide gels (NuPAGE, Invitrogen, cat. no.: NPO321). Western blot analysis was performed using a Semi-Try Blotting Chamber (Biorad) and nitrocellulose membranes (0,2 µm; Schleicher & Schuell) according to the manufacturers protocols. The fusion proteins were detected by using an alkaline phosphatase-conjugated anti-strep tag antibody (IBA, Cat. No.: 2-1503-001) according to the manufacturers protocol.

Generation of the M5Pax8 Cre reporter cell line: The SV40-transformed murine embryonic fibroblast line MEF5/5 (Schwenk et al., Nucl Acids Res 26(6), 1427-32 (1998)) was transfected with the vector pPGKpaX1 (Kellendonk et al., Nucl. Acids Res. 24, 1404-11 (1996)). 10^6 MEF5/5 cells were electroporated with 20 µg pPGKpaX1 plasmid DNA linearised with Sca I and plated into 48-well-plates. The cells were cultured in DMEM/Glutamax medium (Life Technologies) supplemented with 10 % fetal calf serum at 37°C, 10 % CO_2 in humid atmosphere. Two days after transfection the medium was supplemented with 5 µg/ml puromycine (Calbiochem) for the selection of stable integrants. 14 puromycine-resistant clones were expanded and tested by transien transfection with the Cre expression vector pPGK-Cre-pA (SEQ ID NO: 40). In two out of the 14 puromycine-resistant clones, the expression of ß-galactosidase could be detected by staining with X-Gal. One of these clones, M5Pax8, was used as Cre reporter cell line.

Transfection and measurement of ß-galactosidase activity: Fibroblasts (10^6 cells per 24 well plate (Falcon)) were transfected with 25 ng pCMV-I-Cre-pA (see SEQ ID NO:29) or pCMV-I-ß-pA (see SEQ ID NO:30) plasmids using the FuGene transfection reagent (Roche Diagnostics). After 2 days the cells were lysed and the ß-galactosidase activities were determined with the ß-galactosidase reporter gene assay (Roche Diagnostics) according to the manufacturers guidelines using a Lumistar luminometer (MWG).

Histochemical detection of ß-galactosidase activity: To quantitate ß-galactosidase expression, fibroblast cells were washed once with phosphate buffered saline (PBS), and the cells were fixed for 5 minutes at room temperature in a solution of 4% formaldehyde in PBS. Next, the cells were washed twice with PBS and finally incubated in staining solution for 24 hours at 37°C (staining solution: 5 mM K3(Fe(CN)6), 5mM

K4(Fe(CN)6), 2mM MgCl2, 1mg/ml X-Gal (BioMol) in PBS). Blue stained, ß-galactosidase positive cells were detected and distinguished from negative (transparent) cells in a cell culture binocular microscope under 200x magnification. For each determination a minimum of 200 cells was counted.

PCR detection of Cre-mediated recombination: Genomic DNA extracted from tissue samples was subjected to PCR using Taq-polymerase (Gibco BRL Cat. No. 10342-020) using 20 pmol of each primer (sense: 5`-CAT CTC CGG GCC TTT CGA CCT G - 3', antisense: 5' –GCG ATC GGT GCG GGC CTC TTC - 3'; SEQ ID Nos: 41 and 42, respectively). PCR was performed using the following cycle profile: 2' 94°C, 35 x (30" 94°C, 30" 55 °C, 1' 72 °C), 10 min 72 °C. PCR products were separated on a 1,2 % agarose gel.

Example 1

The vector pT7-TACS (SEQ ID NO:16) was constructed for the expression of a TAT-Cre fusion protein in E. coli. The plasmid contains the coding region of the 11 amino acid protein transduction domain of the wild-type HIV TAT protein (Green and Loewenstein, Cell, 55(6):1179-88 (1988); Frankel and Pabo, Cell, 55(6): 1189-93 (1988); SEQ ID NO:10) fused to the N-terminus of Cre recombinase protein sequence. The 10-amino-acid strep tag at the C-terminus allows the detection and purification of the fusion protein using specific antibodies (Schmidt and Skerra, J. Chromatogr A 676: 337-345 (1994)). The protease factor Xa recognition site (Ile-Glu-Gly-Arg) permits the removal of the strep tag by proteolytic cleavage. The estimated molecular weight of the TAT-Cre fusion protein is 42 kDa. A scheme of the TAT-Cre expression vector is depicted in figure 2. For the expression of TAT-Cre, the E. coli strain BL21(DE3)-RIL (Stratagene) was used. This strain carries an IPTG-inducible T7 polymerase gene and additional copies of the tRNA genes for the 'rare

codons' argU, ileY and leuW.

E. coli BL21(DE3)-RIL cells were transformed with pT7-TACS and grown in LB medium containing 100 µg/ml ampicillin. The expression of the 40 kDa TAT-Cre fusion protein could be strongly induced by the addition of 0,5 mM IPTG to the culture medium. Analysis of protein lysates revealed that approximately 50 % of TAT-Cre protein accumulated as insoluble inclusion bodies. The inclusion bodies where extracted and dissolved in 8 M urea. TAT-Cre was subsequently purified from this fraction using ion exchange chromatography. The quantity and purity of TAT-Cre protein was determined using Coomassie stained SDS-PAGE gels and Western blot analysis (figure 3). The purification process yielded TAT-Cre protein extracts of 64 % purity and a concentration of 100 µg/ml.

To analyse the ability of the purified TAT-Cre protein to transduce into cultured cells, we used the fibroblast cell line M5Pax8 (R. Kühn, unpublished) that contains a loxP-containing reporter construct. This reporter, when recombined by Cre recombinase, allows the expression of a ß-galacosidase gene (Buchholz et al, Nucleic Acids Res. 24, 4256-4262, 1996). Cells were cultured for 18 h with increasing concentrations of TAT-Cre protein in serum-free medium and analysed 4 days later for ß-Galacosidase activity. Staining with X-Gal showed that > 50 % of the cells treated with 13,8 µg/ml TAT-Cre protein expressed ß-galactosidase indicating recombination of the loxP-flanked reporter construct had occurred (figure 4). Measurement of ß-galactosidase activity in cell lysates revealed an up to 30-fold higher level of ß-galactosidase activity in comparison to cells which had been transiently transfected with an eukaryotic Cre expression vector (figure 5).

To investigate the activity of TAT-Cre protein in a living organism, we used a transgenic mouse strain carrying a loxP-flanked target for Cre-mediated recombination (Thorey et al., 1998, Mol. Cell. Biol. 18: 3081 – 3088). Mice where treated three times with intraperitoneal injections of 75 µg TAT Cre protein at two-day-intervals and analysed 2 days later. Genomic DNA was

isolated from a variety of organs and subjected to PCR amplification which specifically amplifies a 400 bp fragment of the recombined allele. The deleted allele could be detected in multiple tissues from treated mice indicating TAT-Cre-mediated recombination in these organs (figure 6). This experiments demonstrates that TAT-mediated delivery of active Cre protein works with sufficient efficacy to facilitate inducible gene targeting in cell lines and in living organisms.

Example 2

The vectors pT7-VPCS (SEQ ID NO:17) and pCRT7-ΔVPCS (SEQ ID NO:15) were constructed for the expression of VP22-Cre and ΔVP22-Cre fusion proteins in E. coli. The VP22-Cre gene of pT7-VPCS contains the full length protein translocation domain of the HSV VP22 protein (Elliott and O'Hare, Cell, 88(2): 223-33 (1987), whereas the ΔVP22-Cre gene of pCRT7-ΔVPCS contains a truncated VP22 protein transduction domain (amino acids 159 – 301; Invitrogen; aa 16-157 of SEQ ID NO:14) fused to the N-terminus of Cre recombinase protein sequence. A 10-amino-acid strep tag at the C-terminus of Cre protein sequence allows the detection and purification of the fusion proteins using specific antibodies (Schmidt and Skerra, J. Chromatogr A 676: 337-345 (1994)). The protease factor Xa recognition site permits the removal of the Strep tag by proteolytic cleavage. The estimated molecular weight is 75 kDa for VP22-Cre protein and 60 kDa for ΔVP22-Cre protein. A scheme of the vectors pT7-VPCS and pCRT7-ΔVPCS is depicted in figure 7.

E. coli BL21(DE3)-RIP cells (Stratagene) were transformed with pT7-VPCS and cultured in LB medium containing 100 µg/ml ampicillin. E. coli BL21(DE3)-pLysS cells (Stratagene) were transformed with pCRT7-ΔVPCS and cultured in LB medium containing 25 µg/ml kanamycine and 34 µg/ml chloramphenicol. Expression of the VP22-Cre and ΔVP22-Cre fusion proteins could be induced by the addition of 0,5 mM IPTG to the culture medium. Analysis of protein extracts using Coomassie staining and

Western blotting of SDS-PAGE gels revealed that 50 - 60 % of VP22-Cre and ΔVP22-Cre proteins accumulated as insoluble inclusion bodies. The inclusion bodies where extracted and dissolved in 8 M urea. VP22-Cre and ΔVP22-Cre fusion proteins were subsequently purified using ion exchange chromatography. The quantity and purity of the isolated VP22-Cre and Δ VP22-Cre fusion proteins was determined using Coomassie stained SDS-PAGE gels and Western blot analysis (figure 8).
To analyse the ability of the purified fusion proteins to transduce into cultured cells, we used the fibroblast cell line M5Pax8 that contains a loxP-containing reporter construct. When recombined by Cre recombinase, the reporter allows the expression of a ß-galacosidase gene (Buchholz et al, Nucleic Acids Res. 24, 4256-4262, 1996). The cells where cultured for 18 h with increasing concentrations of VP22-Cre and ΔVP22-Cre in serum-free medium and analysed 4 days later for ß-Galacosidase activity. Staining with X-Gal showed ~2 % blue cells in the cultures treated with up to 15 µg/ml ΔVP22-Cre indicating recombination of the loxP-flanked reporter construct had occurred. In contrast, cell cultures treated with up to 0,5 µg/ml VP22-Cre did not show any X-gal staining (figure 9). Measurement of cell lysates revealed a strong increase of ß-galactosidase activity upon Δ VP22-Cre treatment when compared to untreated cells (figure 10). Genomic DNA was isolated fand subjected to PCR amplification that specifically amplifies a 250 bp fragment of the recombined allele. The deleted allele could be detected in cells treated with both VP22-Cre and Δ VP22-Cre fusion proteins (figure 11).

This experiment demonstrates that VP22-mediated delivery of active Cre protein works with sufficient efficacy to facilitate inducible gene targeting.

SEQUENCE LISTING

<110> ARTEMIS Pharmaceuticals GmbH

<120> Transduction of recombinases for inducible gene
targeting

<130> 010007wo/JH/ml

<140>
<141>

<160> 42

<170> PatentIn Ver. 2.1

<210> 1
<211> 1074
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA sequence
coding for a fusion protein TAT-Cre

<220>
<221> CDS
<222> (1)..(1071)

<400> 1
```
atg ggc tac ggc cgc aag aag cgc cgc caa cgc cgc cgc ggc atg tcc     48
Met Gly Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Gly Met Ser
  1               5                  10                  15

aat tta ctg acc gta cac caa aat ttg cct gca tta ccg gtc gat gca     96
Asn Leu Leu Thr Val His Gln Asn Leu Pro Ala Leu Pro Val Asp Ala
             20                  25                  30

acg agt gat gag gtt cgc aag aac ctg atg gac atg ttc agg gat cgc    144
Thr Ser Asp Glu Val Arg Lys Asn Leu Met Asp Met Phe Arg Asp Arg
         35                  40                  45

cag gcg ttt tct gag cat acc tgg aaa atg ctt ctg tcc gtt tgc cgg    192
Gln Ala Phe Ser Glu His Thr Trp Lys Met Leu Leu Ser Val Cys Arg
     50                  55                  60

tcg tgg gcg gca tgg tgc aag ttg aat aac cgg aaa tgg ttt ccc gca    240
Ser Trp Ala Ala Trp Cys Lys Leu Asn Asn Arg Lys Trp Phe Pro Ala
 65                  70                  75                  80

gaa cct gaa gat gtt cgc gat tat ctt cta tat ctt cag gcg cgc ggt    288
Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu Tyr Leu Gln Ala Arg Gly
                 85                  90                  95

ctg gca gta aaa act atc cag caa cat ttg ggc cag cta aac atg ctt    336
Leu Ala Val Lys Thr Ile Gln Gln His Leu Gly Gln Leu Asn Met Leu
            100                 105                 110

cat cgt cgg tcc ggg ctg cca cga cca agt gac agc aat gct gtt tca    384
His Arg Arg Ser Gly Leu Pro Arg Pro Ser Asp Ser Asn Ala Val Ser
        115                 120                 125
```

```
ctg gtt atg cgg cgg atc cga aaa gaa aac gtt gat gcc ggt gaa cgt     432
Leu Val Met Arg Arg Ile Arg Lys Glu Asn Val Asp Ala Gly Glu Arg
    130                 135                 140

gca aaa cag gct cta gcg ttc gaa cgc act gat ttc gac cag gtt cgt     480
Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Phe Asp Gln Val Arg
145                 150                 155                 160

tca ctc atg gaa aat agc gat cgc tgc cag gat ata cgt aat ctg gca     528
Ser Leu Met Glu Asn Ser Asp Arg Cys Gln Asp Ile Arg Asn Leu Ala
                165                 170                 175

ttt ctg ggg att gct tat aac acc ctg tta cgt ata gcc gaa att gcc     576
Phe Leu Gly Ile Ala Tyr Asn Thr Leu Leu Arg Ile Ala Glu Ile Ala
            180                 185                 190

agg atc agg gtt aaa gat atc tca cgt act gac ggt ggg aga atg tta     624
Arg Ile Arg Val Lys Asp Ile Ser Arg Thr Asp Gly Gly Arg Met Leu
        195                 200                 205

atc cat att ggc aga acg aaa acg ctg gtt agc acc gca ggt gta gag     672
Ile His Ile Gly Arg Thr Lys Thr Leu Val Ser Thr Ala Gly Val Glu
    210                 215                 220

aag gca ctt agc ctg ggg gta act aaa ctg gtc gag cga tgg att tcc     720
Lys Ala Leu Ser Leu Gly Val Thr Lys Leu Val Glu Arg Trp Ile Ser
225                 230                 235                 240

gtc tct ggt gta gct gat gat ccg aat aac tac ctg ttt tgc cgg gtc     768
Val Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys Arg Val
                245                 250                 255

aga aaa aat ggt gtt gcc gcg cca tct gcc acc agc cag cta tca act     816
Arg Lys Asn Gly Val Ala Ala Pro Ser Ala Thr Ser Gln Leu Ser Thr
            260                 265                 270

cgc gcc ctg gaa ggg att ttt gaa gca act cat cga ttg att tac ggc     864
Arg Ala Leu Glu Gly Ile Phe Glu Ala Thr His Arg Leu Ile Tyr Gly
        275                 280                 285

gct aag gat gac tct ggt cag aga tac ctg gcc tgg tct gga cac agt     912
Ala Lys Asp Asp Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly His Ser
    290                 295                 300

gcc cgt gtc gga gcc gcg cga gat atg gcc cgc gct gga gtt tca ata     960
Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val Ser Ile
305                 310                 315                 320

ccg gag atc atg caa gct ggt ggc tgg acc aat gta aat att gtc atg    1008
Pro Glu Ile Met Gln Ala Gly Gly Trp Thr Asn Val Asn Ile Val Met
                325                 330                 335

aac tat atc cgt aac ctg gat agt gaa aca ggg gca atg gtg cgc ctg    1056
Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val Arg Leu
            340                 345                 350

ctg gaa gat ggc gat tag                                            1074
Leu Glu Asp Gly Asp
        355


<210> 2
<211> 357
<212> PRT
```

<213> Artificial Sequence
<223> Description of Artificial Sequence: DNA sequence coding for a fusion protein TAT-Cre

<400> 2
Met Gly Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Gly Met Ser
1 5 10 15

Asn Leu Leu Thr Val His Gln Asn Leu Pro Ala Leu Pro Val Asp Ala
20 25 30

Thr Ser Asp Glu Val Arg Lys Asn Leu Met Asp Met Phe Arg Asp Arg
35 40 45

Gln Ala Phe Ser Glu His Thr Trp Lys Met Leu Leu Ser Val Cys Arg
50 55 60

Ser Trp Ala Ala Trp Cys Lys Leu Asn Asn Arg Lys Trp Phe Pro Ala
65 70 75 80

Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu Tyr Leu Gln Ala Arg Gly
85 90 95

Leu Ala Val Lys Thr Ile Gln Gln His Leu Gly Gln Leu Asn Met Leu
100 105 110

His Arg Arg Ser Gly Leu Pro Arg Pro Ser Asp Ser Asn Ala Val Ser
115 120 125

Leu Val Met Arg Arg Ile Arg Lys Glu Asn Val Asp Ala Gly Glu Arg
130 135 140

Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Phe Asp Gln Val Arg
145 150 155 160

Ser Leu Met Glu Asn Ser Asp Arg Cys Gln Asp Ile Arg Asn Leu Ala
165 170 175

Phe Leu Gly Ile Ala Tyr Asn Thr Leu Leu Arg Ile Ala Glu Ile Ala
180 185 190

Arg Ile Arg Val Lys Asp Ile Ser Arg Thr Asp Gly Gly Arg Met Leu
195 200 205

Ile His Ile Gly Arg Thr Lys Thr Leu Val Ser Thr Ala Gly Val Glu
210 215 220

Lys Ala Leu Ser Leu Gly Val Thr Lys Leu Val Glu Arg Trp Ile Ser
225 230 235 240

Val Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys Arg Val
245 250 255

Arg Lys Asn Gly Val Ala Ala Pro Ser Ala Thr Ser Gln Leu Ser Thr
260 265 270

Arg Ala Leu Glu Gly Ile Phe Glu Ala Thr His Arg Leu Ile Tyr Gly
275 280 285

Ala Lys Asp Asp Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly His Ser
290 295 300

Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val Ser Ile
305 310 315 320

```
Pro Glu Ile Met Gln Ala Gly Gly Trp Thr Asn Val Asn Ile Val Met
                325                 330                 335

Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val Arg Leu
            340                 345                 350

Leu Glu Asp Gly Asp
        355


<210> 3
<211> 1317
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA sequence
      coding for a fusion protein TAT-Flpe

<220>
<221> CDS
<222> (1)..(1311)

<400> 3
atg ggc tac ggc cgc aag aag cgc cgc caa cgc cgc cgc ggc atg agt      48
Met Gly Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Gly Met Ser
  1               5                  10                  15

caa ttt gat ata tta tgt aaa aca cca cct aag gtc ctg gtt cgt cag      96
Gln Phe Asp Ile Leu Cys Lys Thr Pro Pro Lys Val Leu Val Arg Gln
             20                  25                  30

ttt gtg gaa agg ttt gaa aga cct tca ggg gaa aaa ata gca tca tgt     144
Phe Val Glu Arg Phe Glu Arg Pro Ser Gly Glu Lys Ile Ala Ser Cys
         35                  40                  45

gct gct gaa cta acc tat tta tgt tgg atg att act cat aac gga aca     192
Ala Ala Glu Leu Thr Tyr Leu Cys Trp Met Ile Thr His Asn Gly Thr
     50                  55                  60

gca atc aag aga gcc aca ttc atg agc tat aat act atc ata agc aat     240
Ala Ile Lys Arg Ala Thr Phe Met Ser Tyr Asn Thr Ile Ile Ser Asn
 65                  70                  75                  80

tcg ctg agt ttc gat att gtc aac aaa tca ctc cag ttt aaa tac aag     288
Ser Leu Ser Phe Asp Ile Val Asn Lys Ser Leu Gln Phe Lys Tyr Lys
                 85                  90                  95

acg caa aaa gca aca att ctg gaa gcc tca tta aag aaa tta att cct     336
Thr Gln Lys Ala Thr Ile Leu Glu Ala Ser Leu Lys Lys Leu Ile Pro
            100                 105                 110

gct tgg gaa ttt aca att att cct tac aat gga caa aaa cat caa tct     384
Ala Trp Glu Phe Thr Ile Ile Pro Tyr Asn Gly Gln Lys His Gln Ser
        115                 120                 125

gat atc act gat att gta agt agt ttg caa tta cag ttc gaa tca tcg     432
Asp Ile Thr Asp Ile Val Ser Ser Leu Gln Leu Gln Phe Glu Ser Ser
    130                 135                 140

gaa gaa gca gat aag gga aat agc cac agt aaa aaa atg ctt aaa gca     480
Glu Glu Ala Asp Lys Gly Asn Ser His Ser Lys Lys Met Leu Lys Ala
145                 150                 155                 160
```

```
ctt cta agt gag ggt gaa agc atc tgg gag atc act gag aaa ata cta     528
Leu Leu Ser Glu Gly Glu Ser Ile Trp Glu Ile Thr Glu Lys Ile Leu
                165                 170                 175

aat tcg ttt gag tat acc tcg aga ttt aca aaa aca aaa act tta tac     576
Asn Ser Phe Glu Tyr Thr Ser Arg Phe Thr Lys Thr Lys Thr Leu Tyr
            180                 185                 190

caa ttc ctc ttc cta gct act ttc atc aat tgt gga aga ttc agc gat     624
Gln Phe Leu Phe Leu Ala Thr Phe Ile Asn Cys Gly Arg Phe Ser Asp
        195                 200                 205

att aag aac gtt gat ccg aaa tca ttt aaa tta gtc caa aat aag tat     672
Ile Lys Asn Val Asp Pro Lys Ser Phe Lys Leu Val Gln Asn Lys Tyr
    210                 215                 220

ctg gga gta ata atc cag tgt tta gtg aca gag aca aag aca agc gtt     720
Leu Gly Val Ile Ile Gln Cys Leu Val Thr Glu Thr Lys Thr Ser Val
225                 230                 235                 240

agt agg cac ata tac ttc ttt agc gca agg ggt agg atc gat cca ctt     768
Ser Arg His Ile Tyr Phe Phe Ser Ala Arg Gly Arg Ile Asp Pro Leu
                245                 250                 255

gta tat ttg gat gaa ttt ttg agg aat tct gaa cca gtc cta aaa cga     816
Val Tyr Leu Asp Glu Phe Leu Arg Asn Ser Glu Pro Val Leu Lys Arg
            260                 265                 270

gta aat agg acc ggc aat tct tca agc aac aaa cag gaa tac caa tta     864
Val Asn Arg Thr Gly Asn Ser Ser Ser Asn Lys Gln Glu Tyr Gln Leu
        275                 280                 285

tta aaa gat aac tta gtc aga tcg tac aac aag gct ttg aag aaa aat     912
Leu Lys Asp Asn Leu Val Arg Ser Tyr Asn Lys Ala Leu Lys Lys Asn
    290                 295                 300

gcg cct tat cca atc ttt gct ata aag aat ggc cca aaa tct cac att     960
Ala Pro Tyr Pro Ile Phe Ala Ile Lys Asn Gly Pro Lys Ser His Ile
305                 310                 315                 320

gga aga cat ttg atg acc tca ttt ctg tca atg aag ggc cta acg gag    1008
Gly Arg His Leu Met Thr Ser Phe Leu Ser Met Lys Gly Leu Thr Glu
                325                 330                 335

ttg act aat gtt gtg gga aat tgg agc gat aag cgt gct tct gcc gtg    1056
Leu Thr Asn Val Val Gly Asn Trp Ser Asp Lys Arg Ala Ser Ala Val
            340                 345                 350

gcc agg aca acg tat act cat cag ata aca gca ata cct gat cac tac    1104
Ala Arg Thr Thr Tyr Thr His Gln Ile Thr Ala Ile Pro Asp His Tyr
        355                 360                 365

ttc gca cta gtt tct cgg tac tat gca tat gat cca ata tca aag gaa    1152
Phe Ala Leu Val Ser Arg Tyr Tyr Ala Tyr Asp Pro Ile Ser Lys Glu
    370                 375                 380

atg ata gca ttg aag gat gag act aat cca att gag gag tgg cag cat    1200
Met Ile Ala Leu Lys Asp Glu Thr Asn Pro Ile Glu Glu Trp Gln His
385                 390                 395                 400

ata gaa cag cta aag ggt agt gct gaa gga agc ata cga tac ccc gca    1248
Ile Glu Gln Leu Lys Gly Ser Ala Glu Gly Ser Ile Arg Tyr Pro Ala
                405                 410                 415
```

```
tgg aat ggg ata ata tca cag gag gta cta gac tac ctt tca tcc tac    1296
Trp Asn Gly Ile Ile Ser Gln Glu Val Leu Asp Tyr Leu Ser Ser Tyr
            420                 425                 430

ata aat aga cgc ata taatga                                         1317
Ile Asn Arg Arg Ile
        435
```

<210> 4
<211> 437
<212> PRT
<213> Artificial Sequence
<223> Description of Artificial Sequence: DNA sequence coding for a fusion protein TAT-Flpe

<400> 4

```
Met Gly Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Gly Met Ser
  1               5                  10                  15

Gln Phe Asp Ile Leu Cys Lys Thr Pro Pro Lys Val Leu Val Arg Gln
            20                  25                  30

Phe Val Glu Arg Phe Glu Arg Pro Ser Gly Glu Lys Ile Ala Ser Cys
        35                  40                  45

Ala Ala Glu Leu Thr Tyr Leu Cys Trp Met Ile Thr His Asn Gly Thr
    50                  55                  60

Ala Ile Lys Arg Ala Thr Phe Met Ser Tyr Asn Thr Ile Ile Ser Asn
 65                  70                  75                  80

Ser Leu Ser Phe Asp Ile Val Asn Lys Ser Leu Gln Phe Lys Tyr Lys
                 85                  90                  95

Thr Gln Lys Ala Thr Ile Leu Glu Ala Ser Leu Lys Lys Leu Ile Pro
            100                 105                 110

Ala Trp Glu Phe Thr Ile Ile Pro Tyr Asn Gly Gln Lys His Gln Ser
        115                 120                 125

Asp Ile Thr Asp Ile Val Ser Ser Leu Gln Leu Gln Phe Glu Ser Ser
    130                 135                 140

Glu Glu Ala Asp Lys Gly Asn Ser His Ser Lys Lys Met Leu Lys Ala
145                 150                 155                 160

Leu Leu Ser Glu Gly Glu Ser Ile Trp Glu Ile Thr Glu Lys Ile Leu
                165                 170                 175

Asn Ser Phe Glu Tyr Thr Ser Arg Phe Thr Lys Thr Lys Thr Leu Tyr
            180                 185                 190

Gln Phe Leu Phe Leu Ala Thr Phe Ile Asn Cys Gly Arg Phe Ser Asp
        195                 200                 205

Ile Lys Asn Val Asp Pro Lys Ser Phe Lys Leu Val Gln Asn Lys Tyr
    210                 215                 220

Leu Gly Val Ile Ile Gln Cys Leu Val Thr Glu Thr Lys Thr Ser Val
225                 230                 235                 240
```

Ser Arg His Ile Tyr Phe Phe Ser Ala Arg Gly Arg Ile Asp Pro Leu
245 250 255

Val Tyr Leu Asp Glu Phe Leu Arg Asn Ser Glu Pro Val Leu Lys Arg
260 265 270

Val Asn Arg Thr Gly Asn Ser Ser Ser Asn Lys Gln Glu Tyr Gln Leu
275 280 285

Leu Lys Asp Asn Leu Val Arg Ser Tyr Asn Lys Ala Leu Lys Lys Asn
290 295 300

Ala Pro Tyr Pro Ile Phe Ala Ile Lys Asn Gly Pro Lys Ser His Ile
305 310 315 320

Gly Arg His Leu Met Thr Ser Phe Leu Ser Met Lys Gly Leu Thr Glu
325 330 335

Leu Thr Asn Val Val Gly Asn Trp Ser Asp Lys Arg Ala Ser Ala Val
340 345 350

Ala Arg Thr Thr Tyr Thr His Gln Ile Thr Ala Ile Pro Asp His Tyr
355 360 365

Phe Ala Leu Val Ser Arg Tyr Tyr Ala Tyr Asp Pro Ile Ser Lys Glu
370 375 380

Met Ile Ala Leu Lys Asp Glu Thr Asn Pro Ile Glu Glu Trp Gln His
385 390 395 400

Ile Glu Gln Leu Lys Gly Ser Ala Glu Gly Ser Ile Arg Tyr Pro Ala
405 410 415

Trp Asn Gly Ile Ile Ser Gln Glu Val Leu Asp Tyr Leu Ser Ser Tyr
420 425 430

Ile Asn Arg Arg Ile
435

<210> 5
<211> 2004
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA sequence coding for a fusion protein VP22-Cre

<220>
<221> CDS
<222> (1)..(2001)

<400> 5
atg acc tct cgc cgc tcc gtg aag tcg ggt ccg cgg gag gtt ccg cgc 48
Met Thr Ser Arg Arg Ser Val Lys Ser Gly Pro Arg Glu Val Pro Arg
1 5 10 15

gat gag tac gag gat ctg tac tac acc ccg tct tca ggt atg gcg agt 96
Asp Glu Tyr Glu Asp Leu Tyr Tyr Thr Pro Ser Ser Gly Met Ala Ser
20 25 30

```
ccc gat agt ccg cct gac acc tcc cgc cgt ggc gcc cta cag aca cgc    144
Pro Asp Ser Pro Pro Asp Thr Ser Arg Arg Gly Ala Leu Gln Thr Arg
         35                  40                  45

tcg cgc cag agg ggc gag gtc cgt ttc gtc cag tac gac gag tcg gat    192
Ser Arg Gln Arg Gly Glu Val Arg Phe Val Gln Tyr Asp Glu Ser Asp
     50                  55                  60

tat gcc ctc tac ggg ggc tcg tct tcc gaa gac gac gaa cac ccg gag    240
Tyr Ala Leu Tyr Gly Gly Ser Ser Ser Glu Asp Asp Glu His Pro Glu
 65                  70                  75                  80

gtc ccc cgg acg cgg cgt ccc gtt tcc ggg gcg gtt ttg tcc ggc ccg    288
Val Pro Arg Thr Arg Arg Pro Val Ser Gly Ala Val Leu Ser Gly Pro
                 85                  90                  95

ggg cct gcg cgg gcg cct ccg cca ccc gct ggg tcc gga ggg gcc gga    336
Gly Pro Ala Arg Ala Pro Pro Pro Pro Ala Gly Ser Gly Gly Ala Gly
            100                 105                 110

cgc aca ccc acc acc gcc ccc cgg gcc ccc cga acc cag cgg gtg gcg    384
Arg Thr Pro Thr Thr Ala Pro Arg Ala Pro Arg Thr Gln Arg Val Ala
        115                 120                 125

act aag gcc ccc gcg gcc ccg gcg gcg gag acc acc cgc ggc agg aaa    432
Thr Lys Ala Pro Ala Ala Pro Ala Ala Glu Thr Thr Arg Gly Arg Lys
    130                 135                 140

tcg gcc cag cca gaa tcc gcc gca ctc cca gac gcc ccc gcg tcg acg    480
Ser Ala Gln Pro Glu Ser Ala Ala Leu Pro Asp Ala Pro Ala Ser Thr
145                 150                 155                 160

gcg cca acc cga tcc aag aca ccc gcg cag ggg ctg gcc aga aag ctg    528
Ala Pro Thr Arg Ser Lys Thr Pro Ala Gln Gly Leu Ala Arg Lys Leu
                165                 170                 175

cac ttt agc acc gcc ccc cca aac ccc gac gcg cca tgg acc ccc cgg    576
His Phe Ser Thr Ala Pro Pro Asn Pro Asp Ala Pro Trp Thr Pro Arg
            180                 185                 190

gtg gcc ggc ttt aac aag cgc gtc ttc tgc gcc gcg gtc ggg cgc ctg    624
Val Ala Gly Phe Asn Lys Arg Val Phe Cys Ala Ala Val Gly Arg Leu
        195                 200                 205

gcg gcc atg cat gcc cgg atg gcg gcg gtc cag ctc tgg gac atg tcg    672
Ala Ala Met His Ala Arg Met Ala Ala Val Gln Leu Trp Asp Met Ser
    210                 215                 220

cgt ccg cgc aca gac gaa gac ctc aac gaa ctc ctt ggc atc acc acc    720
Arg Pro Arg Thr Asp Glu Asp Leu Asn Glu Leu Leu Gly Ile Thr Thr
225                 230                 235                 240

atc cgc gtg acg gtc tgc gag ggc aaa aac ctg ctt cag cgc gcc aac    768
Ile Arg Val Thr Val Cys Glu Gly Lys Asn Leu Leu Gln Arg Ala Asn
                245                 250                 255

gag ttg gtg aat cca gac gtg gtg cag gac gtc gac gcg gcc acg gcg    816
Glu Leu Val Asn Pro Asp Val Val Gln Asp Val Asp Ala Ala Thr Ala
            260                 265                 270

act cga ggg cgt tct gcg gcg tcg cgc ccc acc gag cga cct cga gcc    864
Thr Arg Gly Arg Ser Ala Ala Ser Arg Pro Thr Glu Arg Pro Arg Ala
        275                 280                 285
```

```
cca gcc cgc tcc gct tct cgc ccc aga cgg ccc gtc gag ggt acc gag    912
Pro Ala Arg Ser Ala Ser Arg Pro Arg Arg Pro Val Glu Gly Thr Glu
    290                 295                 300

ctc gga tcc act agt cca gtg tgg tgg aat tct gca gat atc cag cac    960
Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Gln His
305                 310                 315                 320

agt ggc ggc cgc atg tcc aat tta ctg acc gta cac caa aat ttg cct   1008
Ser Gly Gly Arg Met Ser Asn Leu Leu Thr Val His Gln Asn Leu Pro
                325                 330                 335

gca tta ccg gtc gat gca acg agt gat gag gtt cgc aag aac ctg atg   1056
Ala Leu Pro Val Asp Ala Thr Ser Asp Glu Val Arg Lys Asn Leu Met
            340                 345                 350

gac atg ttc agg gat cgc cag gcg ttt tct gag cat acc tgg aaa atg   1104
Asp Met Phe Arg Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Met
        355                 360                 365

ctt ctg tcc gtt tgc cgg tcg tgg gcg gca tgg tgc aag ttg aat aac   1152
Leu Leu Ser Val Cys Arg Ser Trp Ala Ala Trp Cys Lys Leu Asn Asn
    370                 375                 380

cgg aaa tgg ttt ccc gca gaa cct gaa gat gtt cgc gat tat ctt cta   1200
Arg Lys Trp Phe Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu
385                 390                 395                 400

tat ctt cag gcg cgc ggt ctg gca gta aaa act atc cag caa cat ttg   1248
Tyr Leu Gln Ala Arg Gly Leu Ala Val Lys Thr Ile Gln Gln His Leu
                405                 410                 415

ggc cag cta aac atg ctt cat cgt cgg tcc ggg ctg cca cga cca agt   1296
Gly Gln Leu Asn Met Leu His Arg Arg Ser Gly Leu Pro Arg Pro Ser
            420                 425                 430

gac agc aat gct gtt tca ctg gtt atg cgg cgg atc cga aaa gaa aac   1344
Asp Ser Asn Ala Val Ser Leu Val Met Arg Arg Ile Arg Lys Glu Asn
        435                 440                 445

gtt gat gcc ggt gaa cgt gca aaa cag gct cta gcg ttc gaa cgc act   1392
Val Asp Ala Gly Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr
    450                 455                 460

gat ttc gac cag gtt cgt tca ctc atg gaa aat agc gat cgc tgc cag   1440
Asp Phe Asp Gln Val Arg Ser Leu Met Glu Asn Ser Asp Arg Cys Gln
465                 470                 475                 480

gat ata cgt aat ctg gca ttt ctg ggg att gct tat aac acc ctg tta   1488
Asp Ile Arg Asn Leu Ala Phe Leu Gly Ile Ala Tyr Asn Thr Leu Leu
                485                 490                 495

cgt ata gcc gaa att gcc agg atc agg gtt aaa gat atc tca cgt act   1536
Arg Ile Ala Glu Ile Ala Arg Ile Arg Val Lys Asp Ile Ser Arg Thr
            500                 505                 510

gac ggt ggg aga atg tta atc cat att ggc aga acg aaa acg ctg gtt   1584
Asp Gly Gly Arg Met Leu Ile His Ile Gly Arg Thr Lys Thr Leu Val
        515                 520                 525

agc acc gca ggt gta gag aag gca ctt agc ctg ggg gta act aaa ctg   1632
Ser Thr Ala Gly Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu
    530                 535                 540
```

gtc gag cga tgg att tcc gtc tct ggt gta gct gat gat ccg aat aac 1680
Val Glu Arg Trp Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn
545 550 555 560

tac ctg ttt tgc cgg gtc aga aaa aat ggt gtt gcc gcg cca tct gcc 1728
Tyr Leu Phe Cys Arg Val Arg Lys Asn Gly Val Ala Ala Pro Ser Ala
565 570 575

acc agc cag cta tca act cgc gcc ctg gaa ggg att ttt gaa gca act 1776
Thr Ser Gln Leu Ser Thr Arg Ala Leu Glu Gly Ile Phe Glu Ala Thr
580 585 590

cat cga ttg att tac ggc gct aag gat gac tct ggt cag aga tac ctg 1824
His Arg Leu Ile Tyr Gly Ala Lys Asp Asp Ser Gly Gln Arg Tyr Leu
595 600 605

gcc tgg tct gga cac agt gcc cgt gtc gga gcc gcg cga gat atg gcc 1872
Ala Trp Ser Gly His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala
610 615 620

cgc gct gga gtt tca ata ccg gag atc atg caa gct ggt ggc tgg acc 1920
Arg Ala Gly Val Ser Ile Pro Glu Ile Met Gln Ala Gly Gly Trp Thr
625 630 635 640

aat gta aat att gtc atg aac tat atc cgt aac ctg gat agt gaa aca 1968
Asn Val Asn Ile Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr
645 650 655

ggg gca atg gtg cgc ctg ctg gaa gat ggc gat tag 2004
Gly Ala Met Val Arg Leu Leu Glu Asp Gly Asp
660 665

<210> 6
<211> 667
<212> PRT
<213> Artificial Sequence
<223> Description of Artificial Sequence: DNA sequence
coding for a fusion protein VP22-Cre

<400> 6
Met Thr Ser Arg Arg Ser Val Lys Ser Gly Pro Arg Glu Val Pro Arg
1 5 10 15

Asp Glu Tyr Glu Asp Leu Tyr Tyr Thr Pro Ser Ser Gly Met Ala Ser
20 25 30

Pro Asp Ser Pro Pro Asp Thr Ser Arg Arg Gly Ala Leu Gln Thr Arg
35 40 45

Ser Arg Gln Arg Gly Glu Val Arg Phe Val Gln Tyr Asp Glu Ser Asp
50 55 60

Tyr Ala Leu Tyr Gly Gly Ser Ser Ser Glu Asp Asp Glu His Pro Glu
65 70 75 80

Val Pro Arg Thr Arg Arg Pro Val Ser Gly Ala Val Leu Ser Gly Pro
85 90 95

Gly Pro Ala Arg Ala Pro Pro Pro Pro Ala Gly Ser Gly Gly Ala Gly
100 105 110

Arg Thr Pro Thr Thr Ala Pro Arg Ala Pro Arg Thr Gln Arg Val Ala
115 120 125

Thr Lys Ala Pro Ala Ala Pro Ala Ala Glu Thr Thr Arg Gly Arg Lys
130 135 140

Ser Ala Gln Pro Glu Ser Ala Ala Leu Pro Asp Ala Pro Ala Ser Thr
145 150 155 160

Ala Pro Thr Arg Ser Lys Thr Pro Ala Gln Gly Leu Ala Arg Lys Leu
165 170 175

His Phe Ser Thr Ala Pro Pro Asn Pro Asp Ala Pro Trp Thr Pro Arg
180 185 190

Val Ala Gly Phe Asn Lys Arg Val Phe Cys Ala Ala Val Gly Arg Leu
195 200 205

Ala Ala Met His Ala Arg Met Ala Ala Val Gln Leu Trp Asp Met Ser
210 215 220

Arg Pro Arg Thr Asp Glu Asp Leu Asn Glu Leu Leu Gly Ile Thr Thr
225 230 235 240

Ile Arg Val Thr Val Cys Glu Gly Lys Asn Leu Leu Gln Arg Ala Asn
245 250 255

Glu Leu Val Asn Pro Asp Val Val Gln Asp Val Asp Ala Ala Thr Ala
260 265 270

Thr Arg Gly Arg Ser Ala Ala Ser Arg Pro Thr Glu Arg Pro Arg Ala
275 280 285

Pro Ala Arg Ser Ala Ser Arg Pro Arg Arg Pro Val Glu Gly Thr Glu
290 295 300

Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Gln His
305 310 315 320

Ser Gly Gly Arg Met Ser Asn Leu Leu Thr Val His Gln Asn Leu Pro
325 330 335

Ala Leu Pro Val Asp Ala Thr Ser Asp Glu Val Arg Lys Asn Leu Met
340 345 350

Asp Met Phe Arg Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Met
355 360 365

Leu Leu Ser Val Cys Arg Ser Trp Ala Ala Trp Cys Lys Leu Asn Asn
370 375 380

Arg Lys Trp Phe Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu
385 390 395 400

Tyr Leu Gln Ala Arg Gly Leu Ala Val Lys Thr Ile Gln Gln His Leu
405 410 415

Gly Gln Leu Asn Met Leu His Arg Arg Ser Gly Leu Pro Arg Pro Ser
420 425 430

Asp Ser Asn Ala Val Ser Leu Val Met Arg Arg Ile Arg Lys Glu Asn
435 440 445

Val Asp Ala Gly Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr
450 455 460

Asp Phe Asp Gln Val Arg Ser Leu Met Glu Asn Ser Asp Arg Cys Gln
465 470 475 480

Asp Ile Arg Asn Leu Ala Phe Leu Gly Ile Ala Tyr Asn Thr Leu Leu
485 490 495

Arg Ile Ala Glu Ile Ala Arg Ile Arg Val Lys Asp Ile Ser Arg Thr
500 505 510

Asp Gly Gly Arg Met Leu Ile His Ile Gly Arg Thr Lys Thr Leu Val
515 520 525

Ser Thr Ala Gly Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu
530 535 540

Val Glu Arg Trp Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn
545 550 555 560

Tyr Leu Phe Cys Arg Val Arg Lys Asn Gly Val Ala Ala Pro Ser Ala
565 570 575

Thr Ser Gln Leu Ser Thr Arg Ala Leu Glu Gly Ile Phe Glu Ala Thr
580 585 590

His Arg Leu Ile Tyr Gly Ala Lys Asp Asp Ser Gly Gln Arg Tyr Leu
595 600 605

Ala Trp Ser Gly His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala
610 615 620

Arg Ala Gly Val Ser Ile Pro Glu Ile Met Gln Ala Gly Gly Trp Thr
625 630 635 640

Asn Val Asn Ile Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr
645 650 655

Gly Ala Met Val Arg Leu Leu Glu Asp Gly Asp
660 665

<210> 7
<211> 2247
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA sequence coding for a fusion protein VP22-Flpe

<220>
<221> CDS
<222> (1)..(2241)

<400> 7
atg acc tct cgc cgc tcc gtg aag tcg ggt ccg cgg gag gtt ccg cgc 48
Met Thr Ser Arg Arg Ser Val Lys Ser Gly Pro Arg Glu Val Pro Arg
1 5 10 15

gat gag tac gag gat ctg tac tac acc ccg tct tca ggt atg gcg agt 96
Asp Glu Tyr Glu Asp Leu Tyr Tyr Thr Pro Ser Ser Gly Met Ala Ser
20 25 30

```
ccc gat agt ccg cct gac acc tcc cgc cgt ggc gcc cta cag aca cgc      144
Pro Asp Ser Pro Pro Asp Thr Ser Arg Arg Gly Ala Leu Gln Thr Arg
         35                  40                  45

tcg cgc cag agg ggc gag gtc cgt ttc gtc cag tac gac gag tcg gat      192
Ser Arg Gln Arg Gly Glu Val Arg Phe Val Gln Tyr Asp Glu Ser Asp
     50                  55                  60

tat gcc ctc tac ggg ggc tcg tct tcc gaa gac gac gaa cac ccg gag      240
Tyr Ala Leu Tyr Gly Gly Ser Ser Ser Glu Asp Asp Glu His Pro Glu
 65                  70                  75                  80

gtc ccc cgg acg cgg cgt ccc gtt tcc ggg gcg gtt ttg tcc ggc ccg      288
Val Pro Arg Thr Arg Arg Pro Val Ser Gly Ala Val Leu Ser Gly Pro
                 85                  90                  95

ggg cct gcg cgg gcg cct ccg cca ccc gct ggg tcc gga ggg gcc gga      336
Gly Pro Ala Arg Ala Pro Pro Pro Pro Ala Gly Ser Gly Gly Ala Gly
            100                 105                 110

cgc aca ccc acc acc gcc ccc cgg gcc ccc cga acc cag cgg gtg gcg      384
Arg Thr Pro Thr Thr Ala Pro Arg Ala Pro Arg Thr Gln Arg Val Ala
        115                 120                 125

act aag gcc ccc gcg gcc ccg gcg gcg gag acc acc cgc ggc agg aaa      432
Thr Lys Ala Pro Ala Ala Pro Ala Ala Glu Thr Thr Arg Gly Arg Lys
    130                 135                 140

tcg gcc cag cca gaa tcc gcc gca ctc cca gac gcc ccc gcg tcg acg      480
Ser Ala Gln Pro Glu Ser Ala Ala Leu Pro Asp Ala Pro Ala Ser Thr
145                 150                 155                 160

gcg cca acc cga tcc aag aca ccc gcg cag ggg ctg gcc aga aag ctg      528
Ala Pro Thr Arg Ser Lys Thr Pro Ala Gln Gly Leu Ala Arg Lys Leu
                165                 170                 175

cac ttt agc acc gcc ccc cca aac ccc gac gcg cca tgg acc ccc cgg      576
His Phe Ser Thr Ala Pro Pro Asn Pro Asp Ala Pro Trp Thr Pro Arg
            180                 185                 190

gtg gcc ggc ttt aac aag cgc gtc ttc tgc gcc gcg gtc ggg cgc ctg      624
Val Ala Gly Phe Asn Lys Arg Val Phe Cys Ala Ala Val Gly Arg Leu
        195                 200                 205

gcg gcc atg cat gcc cgg atg gcg gcg gtc cag ctc tgg gac atg tcg      672
Ala Ala Met His Ala Arg Met Ala Ala Val Gln Leu Trp Asp Met Ser
    210                 215                 220

cgt ccg cgc aca gac gaa gac ctc aac gaa ctc ctt ggc atc acc acc      720
Arg Pro Arg Thr Asp Glu Asp Leu Asn Glu Leu Leu Gly Ile Thr Thr
225                 230                 235                 240

atc cgc gtg acg gtc tgc gag ggc aaa aac ctg ctt cag cgc gcc aac      768
Ile Arg Val Thr Val Cys Glu Gly Lys Asn Leu Leu Gln Arg Ala Asn
                245                 250                 255

gag ttg gtg aat cca gac gtg gtg cag gac gtc gac gcg gcc acg gcg      816
Glu Leu Val Asn Pro Asp Val Val Gln Asp Val Asp Ala Ala Thr Ala
            260                 265                 270

act cga ggg cgt tct gcg gcg tcg cgc ccc acc gag cga cct cga gcc      864
Thr Arg Gly Arg Ser Ala Ala Ser Arg Pro Thr Glu Arg Pro Arg Ala
        275                 280                 285
```

cca gcc cgc tcc gct tct cgc ccc aga cgg ccc gtc gag ggt acc gag 912
Pro Ala Arg Ser Ala Ser Arg Pro Arg Arg Pro Val Glu Gly Thr Glu
290 295 300

ctc gga tcc act agt cca gtg tgg tgg aat tct gca gat atc cag cac 960
Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Gln His
305 310 315 320

agt ggc ggc cgc atg agt caa ttt gat ata tta tgt aaa aca cca cct 1008
Ser Gly Gly Arg Met Ser Gln Phe Asp Ile Leu Cys Lys Thr Pro Pro
325 330 335

aag gtc ctg gtt cgt cag ttt gtg gaa agg ttt gaa aga cct tca ggg 1056
Lys Val Leu Val Arg Gln Phe Val Glu Arg Phe Glu Arg Pro Ser Gly
340 345 350

gaa aaa ata gca tca tgt gct gct gaa cta acc tat tta tgt tgg atg 1104
Glu Lys Ile Ala Ser Cys Ala Ala Glu Leu Thr Tyr Leu Cys Trp Met
355 360 365

att act cat aac gga aca gca atc aag aga gcc aca ttc atg agc tat 1152
Ile Thr His Asn Gly Thr Ala Ile Lys Arg Ala Thr Phe Met Ser Tyr
370 375 380

aat act atc ata agc aat tcg ctg agt ttc gat att gtc aac aaa tca 1200
Asn Thr Ile Ile Ser Asn Ser Leu Ser Phe Asp Ile Val Asn Lys Ser
385 390 395 400

ctc cag ttt aaa tac aag acg caa aaa gca aca att ctg gaa gcc tca 1248
Leu Gln Phe Lys Tyr Lys Thr Gln Lys Ala Thr Ile Leu Glu Ala Ser
405 410 415

tta aag aaa tta att cct gct tgg gaa ttt aca att att cct tac aat 1296
Leu Lys Lys Leu Ile Pro Ala Trp Glu Phe Thr Ile Ile Pro Tyr Asn
420 425 430

gga caa aaa cat caa tct gat atc act gat att gta agt agt ttg caa 1344
Gly Gln Lys His Gln Ser Asp Ile Thr Asp Ile Val Ser Ser Leu Gln
435 440 445

tta cag ttc gaa tca tcg gaa gaa gca gat aag gga aat agc cac agt 1392
Leu Gln Phe Glu Ser Ser Glu Glu Ala Asp Lys Gly Asn Ser His Ser
450 455 460

aaa aaa atg ctt aaa gca ctt cta agt gag ggt gaa agc atc tgg gag 1440
Lys Lys Met Leu Lys Ala Leu Leu Ser Glu Gly Glu Ser Ile Trp Glu
465 470 475 480

atc act gag aaa ata cta aat tcg ttt gag tat acc tcg aga ttt aca 1488
Ile Thr Glu Lys Ile Leu Asn Ser Phe Glu Tyr Thr Ser Arg Phe Thr
485 490 495

aaa aca aaa act tta tac caa ttc ctc ttc cta gct act ttc atc aat 1536
Lys Thr Lys Thr Leu Tyr Gln Phe Leu Phe Leu Ala Thr Phe Ile Asn
500 505 510

tgt gga aga ttc agc gat att aag aac gtt gat ccg aaa tca ttt aaa 1584
Cys Gly Arg Phe Ser Asp Ile Lys Asn Val Asp Pro Lys Ser Phe Lys
515 520 525

tta gtc caa aat aag tat ctg gga gta ata atc cag tgt tta gtg aca 1632
Leu Val Gln Asn Lys Tyr Leu Gly Val Ile Ile Gln Cys Leu Val Thr
530 535 540

```
gag aca aag aca agc gtt agt agg cac ata tac ttc ttt agc gca agg    1680
Glu Thr Lys Thr Ser Val Ser Arg His Ile Tyr Phe Phe Ser Ala Arg
545                 550                 555                 560

ggt agg atc gat cca ctt gta tat ttg gat gaa ttt ttg agg aat tct    1728
Gly Arg Ile Asp Pro Leu Val Tyr Leu Asp Glu Phe Leu Arg Asn Ser
                565                 570                 575

gaa cca gtc cta aaa cga gta aat agg acc ggc aat tct tca agc aac    1776
Glu Pro Val Leu Lys Arg Val Asn Arg Thr Gly Asn Ser Ser Ser Asn
            580                 585                 590

aaa cag gaa tac caa tta tta aaa gat aac tta gtc aga tcg tac aac    1824
Lys Gln Glu Tyr Gln Leu Leu Lys Asp Asn Leu Val Arg Ser Tyr Asn
        595                 600                 605

aag gct ttg aag aaa aat gcg cct tat cca atc ttt gct ata aag aat    1872
Lys Ala Leu Lys Lys Asn Ala Pro Tyr Pro Ile Phe Ala Ile Lys Asn
    610                 615                 620

ggc cca aaa tct cac att gga aga cat ttg atg acc tca ttt ctg tca    1920
Gly Pro Lys Ser His Ile Gly Arg His Leu Met Thr Ser Phe Leu Ser
625                 630                 635                 640

atg aag ggc cta acg gag ttg act aat gtt gtg gga aat tgg agc gat    1968
Met Lys Gly Leu Thr Glu Leu Thr Asn Val Val Gly Asn Trp Ser Asp
                645                 650                 655

aag cgt gct tct gcc gtg gcc agg aca acg tat act cat cag ata aca    2016
Lys Arg Ala Ser Ala Val Ala Arg Thr Thr Tyr Thr His Gln Ile Thr
            660                 665                 670

gca ata cct gat cac tac ttc gca cta gtt tct cgg tac tat gca tat    2064
Ala Ile Pro Asp His Tyr Phe Ala Leu Val Ser Arg Tyr Tyr Ala Tyr
        675                 680                 685

gat cca ata tca aag gaa atg ata gca ttg aag gat gag act aat cca    2112
Asp Pro Ile Ser Lys Glu Met Ile Ala Leu Lys Asp Glu Thr Asn Pro
    690                 695                 700

att gag gag tgg cag cat ata gaa cag cta aag ggt agt gct gaa gga    2160
Ile Glu Glu Trp Gln His Ile Glu Gln Leu Lys Gly Ser Ala Glu Gly
705                 710                 715                 720

agc ata cga tac ccc gca tgg aat ggg ata ata tca cag gag gta cta    2208
Ser Ile Arg Tyr Pro Ala Trp Asn Gly Ile Ile Ser Gln Glu Val Leu
                725                 730                 735

gac tac ctt tca tcc tac ata aat aga cgc ata taatga                 2247
Asp Tyr Leu Ser Ser Tyr Ile Asn Arg Arg Ile
            740                 745
```

<210> 8
<211> 747
<212> PRT
<213> Artificial Sequence
<223> Description of Artificial Sequence: DNA sequence coding for a fusion protein VP22-Flpe

<400> 8

```
Met Thr Ser Arg Arg Ser Val Lys Ser Gly Pro Arg Glu Val Pro Arg
  1               5                  10                  15
```

```
Asp Glu Tyr Glu Asp Leu Tyr Tyr Thr Pro Ser Ser Gly Met Ala Ser
            20                  25                  30

Pro Asp Ser Pro Pro Asp Thr Ser Arg Arg Gly Ala Leu Gln Thr Arg
        35                  40                  45

Ser Arg Gln Arg Gly Glu Val Arg Phe Val Gln Tyr Asp Glu Ser Asp
    50                  55                  60

Tyr Ala Leu Tyr Gly Gly Ser Ser Ser Glu Asp Asp Glu His Pro Glu
65                  70                  75                  80

Val Pro Arg Thr Arg Arg Pro Val Ser Gly Ala Val Leu Ser Gly Pro
                85                  90                  95

Gly Pro Ala Arg Ala Pro Pro Pro Pro Ala Gly Ser Gly Gly Ala Gly
            100                 105                 110

Arg Thr Pro Thr Thr Ala Pro Arg Ala Pro Arg Thr Gln Arg Val Ala
        115                 120                 125

Thr Lys Ala Pro Ala Ala Pro Ala Ala Glu Thr Thr Arg Gly Arg Lys
    130                 135                 140

Ser Ala Gln Pro Glu Ser Ala Ala Leu Pro Asp Ala Pro Ala Ser Thr
145                 150                 155                 160

Ala Pro Thr Arg Ser Lys Thr Pro Ala Gln Gly Leu Ala Arg Lys Leu
                165                 170                 175

His Phe Ser Thr Ala Pro Pro Asn Pro Asp Ala Pro Trp Thr Pro Arg
            180                 185                 190

Val Ala Gly Phe Asn Lys Arg Val Phe Cys Ala Ala Val Gly Arg Leu
        195                 200                 205

Ala Ala Met His Ala Arg Met Ala Ala Val Gln Leu Trp Asp Met Ser
    210                 215                 220

Arg Pro Arg Thr Asp Glu Asp Leu Asn Glu Leu Leu Gly Ile Thr Thr
225                 230                 235                 240

Ile Arg Val Thr Val Cys Glu Gly Lys Asn Leu Leu Gln Arg Ala Asn
                245                 250                 255

Glu Leu Val Asn Pro Asp Val Val Gln Asp Val Asp Ala Ala Thr Ala
            260                 265                 270

Thr Arg Gly Arg Ser Ala Ala Ser Arg Pro Thr Glu Arg Pro Arg Ala
        275                 280                 285

Pro Ala Arg Ser Ala Ser Arg Pro Arg Arg Pro Val Glu Gly Thr Glu
    290                 295                 300

Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Gln His
305                 310                 315                 320

Ser Gly Gly Arg Met Ser Gln Phe Asp Ile Leu Cys Lys Thr Pro Pro
                325                 330                 335

Lys Val Leu Val Arg Gln Phe Val Glu Arg Phe Glu Arg Pro Ser Gly
            340                 345                 350
```

```
Glu Lys Ile Ala Ser Cys Ala Ala Glu Leu Thr Tyr Leu Cys Trp Met
        355                 360                 365

Ile Thr His Asn Gly Thr Ala Ile Lys Arg Ala Thr Phe Met Ser Tyr
    370                 375                 380

Asn Thr Ile Ile Ser Asn Ser Leu Ser Phe Asp Ile Val Asn Lys Ser
385                 390                 395                 400

Leu Gln Phe Lys Tyr Lys Thr Gln Lys Ala Thr Ile Leu Glu Ala Ser
                405                 410                 415

Leu Lys Lys Leu Ile Pro Ala Trp Glu Phe Thr Ile Ile Pro Tyr Asn
            420                 425                 430

Gly Gln Lys His Gln Ser Asp Ile Thr Asp Ile Val Ser Ser Leu Gln
        435                 440                 445

Leu Gln Phe Glu Ser Ser Glu Glu Ala Asp Lys Gly Asn Ser His Ser
    450                 455                 460

Lys Lys Met Leu Lys Ala Leu Leu Ser Glu Gly Glu Ser Ile Trp Glu
465                 470                 475                 480

Ile Thr Glu Lys Ile Leu Asn Ser Phe Glu Tyr Thr Ser Arg Phe Thr
                485                 490                 495

Lys Thr Lys Thr Leu Tyr Gln Phe Leu Phe Leu Ala Thr Phe Ile Asn
            500                 505                 510

Cys Gly Arg Phe Ser Asp Ile Lys Asn Val Asp Pro Lys Ser Phe Lys
        515                 520                 525

Leu Val Gln Asn Lys Tyr Leu Gly Val Ile Ile Gln Cys Leu Val Thr
    530                 535                 540

Glu Thr Lys Thr Ser Val Ser Arg His Ile Tyr Phe Phe Ser Ala Arg
545                 550                 555                 560

Gly Arg Ile Asp Pro Leu Val Tyr Leu Asp Glu Phe Leu Arg Asn Ser
                565                 570                 575

Glu Pro Val Leu Lys Arg Val Asn Arg Thr Gly Asn Ser Ser Ser Asn
            580                 585                 590

Lys Gln Glu Tyr Gln Leu Leu Lys Asp Asn Leu Val Arg Ser Tyr Asn
        595                 600                 605

Lys Ala Leu Lys Lys Asn Ala Pro Tyr Pro Ile Phe Ala Ile Lys Asn
    610                 615                 620

Gly Pro Lys Ser His Ile Gly Arg His Leu Met Thr Ser Phe Leu Ser
625                 630                 635                 640

Met Lys Gly Leu Thr Glu Leu Thr Asn Val Val Gly Asn Trp Ser Asp
                645                 650                 655

Lys Arg Ala Ser Ala Val Ala Arg Thr Thr Tyr Thr His Gln Ile Thr
            660                 665                 670

Ala Ile Pro Asp His Tyr Phe Ala Leu Val Ser Arg Tyr Tyr Ala Tyr
        675                 680                 685
```

```
Asp Pro Ile Ser Lys Glu Met Ile Ala Leu Lys Asp Glu Thr Asn Pro
    690                 695                 700

Ile Glu Glu Trp Gln His Ile Glu Gln Leu Lys Gly Ser Ala Glu Gly
705                 710                 715                 720

Ser Ile Arg Tyr Pro Ala Trp Asn Gly Ile Ile Ser Gln Glu Val Leu
                725                 730                 735

Asp Tyr Leu Ser Ser Tyr Ile Asn Arg Arg Ile
            740                 745


<210> 9
<211> 33
<212> DNA
<213> Human immunodeficiency virus

<400> 9
tacggccgca agaagcgccg ccaacgccgc cgc                                33


<210> 10
<211> 11
<212> PRT
<213> Human immunodeficiency virus

<400> 10
Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg
  1               5                  10


<210> 11
<211> 42
<212> DNA
<213> Human immunodeficiency virus

<220>
<221> CDS
<222> (4)..(42)

<400> 11
atg ggc tac ggc cgc aag aag cgc cgc caa cgc cgc cgc ggc             42
    Gly Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Gly
      1               5                  10


<210> 12
<211> 13
<212> PRT
<213> Human immunodeficiency virus

<400> 12
Gly Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Gly
  1               5                  10


<210> 13
<211> 1623
<212> DNA
<213> Artificial Sequence

<220>
```

```
<223> Description of Artificial Sequence: DNA sequence
      coding for a fusion protein deltaVP22cre-StrepTag

<220>
<221> CDS
<222> (1)..(1617)

<400> 13
atg gct agc atg act ggt gga cag caa atg ggt cgg gat ccg tcg acg      48
Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Asp Pro Ser Thr
  1               5                  10                  15

gcg cca acc cga tcc aag aca ccc gcg cag ggg ctg gcc aga aag ctg      96
Ala Pro Thr Arg Ser Lys Thr Pro Ala Gln Gly Leu Ala Arg Lys Leu
             20                  25                  30

cac ttt agc acc gcc ccc cca aac ccc gac gcg cca tgg acc ccc cgg     144
His Phe Ser Thr Ala Pro Pro Asn Pro Asp Ala Pro Trp Thr Pro Arg
         35                  40                  45

gtg gcc ggc ttt aac aag cgc gtc ttc tgc gcc gcg gtc ggg cgc ctg     192
Val Ala Gly Phe Asn Lys Arg Val Phe Cys Ala Ala Val Gly Arg Leu
     50                  55                  60

gcg gcc atg cat gcc cgg atg gcg gct gtc cag ctc tgg gac atg tcg     240
Ala Ala Met His Ala Arg Met Ala Ala Val Gln Leu Trp Asp Met Ser
 65                  70                  75                  80

cgt ccg cgc aca gac gaa gac ctc aac gaa ctc ctt ggc atc acc acc     288
Arg Pro Arg Thr Asp Glu Asp Leu Asn Glu Leu Leu Gly Ile Thr Thr
                 85                  90                  95

atc cgc gtg acg gtc tgc gag ggc aaa aac ctg ctt cag cgc gcc aac     336
Ile Arg Val Thr Val Cys Glu Gly Lys Asn Leu Leu Gln Arg Ala Asn
            100                 105                 110

gag ttg gtg aat cca gac gtg gtg cag gac gtc gac gcg gcc acg gcg     384
Glu Leu Val Asn Pro Asp Val Val Gln Asp Val Asp Ala Ala Thr Ala
        115                 120                 125

act cga ggg cgt tct gcg gcg tcg cgc ccc acc gag cga cct cga gcc     432
Thr Arg Gly Arg Ser Ala Ala Ser Arg Pro Thr Glu Arg Pro Arg Ala
    130                 135                 140

cca gcc cgc tcc gct tct cgc ccc aga cgg ccc gtc gag ggt acc gag     480
Pro Ala Arg Ser Ala Ser Arg Pro Arg Arg Pro Val Glu Gly Thr Glu
145                 150                 155                 160

ctc gga tcc act agt cca gtg tgg tgg aat tct gca gat atc cag cac     528
Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Gln His
                165                 170                 175

agt ggc ggc cgc atg tcc aat tta ctg acc gta cac caa aat ttg cct     576
Ser Gly Gly Arg Met Ser Asn Leu Leu Thr Val His Gln Asn Leu Pro
            180                 185                 190

gca tta ccg gtc gat gca acg agt gat gag gtt cgc aag aac ctg atg     624
Ala Leu Pro Val Asp Ala Thr Ser Asp Glu Val Arg Lys Asn Leu Met
        195                 200                 205

gac atg ttc agg gat cgc cag gcg ttt tct gag cat acc tgg aaa atg     672
Asp Met Phe Arg Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Met
    210                 215                 220
```

```
ctt ctg tcc gtt tgc cgg tcg tgg gcg gca tgg tgc aag ttg aat aac    720
Leu Leu Ser Val Cys Arg Ser Trp Ala Ala Trp Cys Lys Leu Asn Asn
225                 230                 235                 240

cgg aaa tgg ttt ccc gca gaa cct gaa gat gtt cgc gat tat ctt cta    768
Arg Lys Trp Phe Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu
                245                 250                 255

tat ctt cag gcg cgc ggt ctg gca gta aaa act atc cag caa cat ttg    816
Tyr Leu Gln Ala Arg Gly Leu Ala Val Lys Thr Ile Gln Gln His Leu
            260                 265                 270

ggc cag cta aac atg ctt cat cgt cgg tcc ggg ctg cca cga cca agt    864
Gly Gln Leu Asn Met Leu His Arg Arg Ser Gly Leu Pro Arg Pro Ser
        275                 280                 285

gac agc aat gct gtt tca ctg gtt atg cgg cgg atc cga aaa gaa aac    912
Asp Ser Asn Ala Val Ser Leu Val Met Arg Arg Ile Arg Lys Glu Asn
    290                 295                 300

gtt gat gcc ggt gaa cgt gca aaa cag gct cta gcg ttc gaa cgc act    960
Val Asp Ala Gly Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr
305                 310                 315                 320

gat ttc gac cag gtt cgt tca ctc atg gaa aat agc gat cgc tgc cag    1008
Asp Phe Asp Gln Val Arg Ser Leu Met Glu Asn Ser Asp Arg Cys Gln
                325                 330                 335

gat ata cgt aat ctg gca ttt ctg ggg att gct tat aac acc ctg tta    1056
Asp Ile Arg Asn Leu Ala Phe Leu Gly Ile Ala Tyr Asn Thr Leu Leu
            340                 345                 350

cgt ata gcc gaa att gcc agg atc agg gtt aaa gat atc tca cgt act    1104
Arg Ile Ala Glu Ile Ala Arg Ile Arg Val Lys Asp Ile Ser Arg Thr
        355                 360                 365

gac ggt ggg aga atg tta atc cat att ggc aga acg aaa acg ctg gtt    1152
Asp Gly Gly Arg Met Leu Ile His Ile Gly Arg Thr Lys Thr Leu Val
    370                 375                 380

agc acc gca ggt gta gag aag gca ctt agc ctg ggg gta act aaa ctg    1200
Ser Thr Ala Gly Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu
385                 390                 395                 400

gtc gag cga tgg att tcc gtc tct ggt gta gct gat gat ccg aat aac    1248
Val Glu Arg Trp Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn
                405                 410                 415

tac ctg ttt tgc cgg gtc aga aaa aat ggt gtt gcc gcg cca tct gcc    1296
Tyr Leu Phe Cys Arg Val Arg Lys Asn Gly Val Ala Ala Pro Ser Ala
            420                 425                 430

acc agc cag cta tca act cgc gcc ctg gaa ggg att ttt gaa gca act    1344
Thr Ser Gln Leu Ser Thr Arg Ala Leu Glu Gly Ile Phe Glu Ala Thr
        435                 440                 445

cat cga ttg att tac ggc gct aag gat gac tct ggt cag aga tac ctg    1392
His Arg Leu Ile Tyr Gly Ala Lys Asp Asp Ser Gly Gln Arg Tyr Leu
    450                 455                 460

gcc tgg tct gga cac agt gcc cgt gtc gga gcc gcg cga gat atg gcc    1440
Ala Trp Ser Gly His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala
465                 470                 475                 480
```

```
cgc gct gga gtt tca ata ccg gag atc atg caa gct ggt ggc tgg acc    1488
Arg Ala Gly Val Ser Ile Pro Glu Ile Met Gln Ala Gly Gly Trp Thr
                485                 490                 495

aat gta aat att gtc atg aac tat atc cgt aac ctg gat agt gaa aca    1536
Asn Val Asn Ile Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr
            500                 505                 510

ggg gca atg gtg cgc ctg ctg gaa gat ggc gat ggt atc gaa ggt cgt    1584
Gly Ala Met Val Arg Leu Leu Glu Asp Gly Asp Gly Ile Glu Gly Arg
        515                 520                 525

ggt agc gct tgg cgt cac ccg cag ttc ggt ggt taataa                 1623
Gly Ser Ala Trp Arg His Pro Gln Phe Gly Gly
    530                 535


<210> 14
<211> 539
<212> PRT
<213> Artificial Sequence
<223> Description of Artificial Sequence: DNA sequence
      coding for a fusion protein deltaVP22cre-StrepTag

<400> 14
Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Asp Pro Ser Thr
  1               5                  10                  15

Ala Pro Thr Arg Ser Lys Thr Pro Ala Gln Gly Leu Ala Arg Lys Leu
             20                  25                  30

His Phe Ser Thr Ala Pro Pro Asn Pro Asp Ala Pro Trp Thr Pro Arg
         35                  40                  45

Val Ala Gly Phe Asn Lys Arg Val Phe Cys Ala Ala Val Gly Arg Leu
     50                  55                  60

Ala Ala Met His Ala Arg Met Ala Ala Val Gln Leu Trp Asp Met Ser
 65                  70                  75                  80

Arg Pro Arg Thr Asp Glu Asp Leu Asn Glu Leu Leu Gly Ile Thr Thr
                 85                  90                  95

Ile Arg Val Thr Val Cys Glu Gly Lys Asn Leu Leu Gln Arg Ala Asn
            100                 105                 110

Glu Leu Val Asn Pro Asp Val Val Gln Asp Val Asp Ala Ala Thr Ala
        115                 120                 125

Thr Arg Gly Arg Ser Ala Ala Ser Arg Pro Thr Glu Arg Pro Arg Ala
    130                 135                 140

Pro Ala Arg Ser Ala Ser Arg Pro Arg Arg Pro Val Glu Gly Thr Glu
145                 150                 155                 160

Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Gln His
                165                 170                 175

Ser Gly Gly Arg Met Ser Asn Leu Leu Thr Val His Gln Asn Leu Pro
            180                 185                 190

Ala Leu Pro Val Asp Ala Thr Ser Asp Glu Val Arg Lys Asn Leu Met
        195                 200                 205
```

```
Asp Met Phe Arg Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Met
    210                 215                 220

Leu Leu Ser Val Cys Arg Ser Trp Ala Ala Trp Cys Lys Leu Asn Asn
225                 230                 235                 240

Arg Lys Trp Phe Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu
                245                 250                 255

Tyr Leu Gln Ala Arg Gly Leu Ala Val Lys Thr Ile Gln Gln His Leu
            260                 265                 270

Gly Gln Leu Asn Met Leu His Arg Arg Ser Gly Leu Pro Arg Pro Ser
        275                 280                 285

Asp Ser Asn Ala Val Ser Leu Val Met Arg Arg Ile Arg Lys Glu Asn
    290                 295                 300

Val Asp Ala Gly Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr
305                 310                 315                 320

Asp Phe Asp Gln Val Arg Ser Leu Met Glu Asn Ser Asp Arg Cys Gln
                325                 330                 335

Asp Ile Arg Asn Leu Ala Phe Leu Gly Ile Ala Tyr Asn Thr Leu Leu
            340                 345                 350

Arg Ile Ala Glu Ile Ala Arg Ile Arg Val Lys Asp Ile Ser Arg Thr
        355                 360                 365

Asp Gly Gly Arg Met Leu Ile His Ile Gly Arg Thr Lys Thr Leu Val
    370                 375                 380

Ser Thr Ala Gly Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu
385                 390                 395                 400

Val Glu Arg Trp Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn
                405                 410                 415

Tyr Leu Phe Cys Arg Val Arg Lys Asn Gly Val Ala Ala Pro Ser Ala
            420                 425                 430

Thr Ser Gln Leu Ser Thr Arg Ala Leu Glu Gly Ile Phe Glu Ala Thr
        435                 440                 445

His Arg Leu Ile Tyr Gly Ala Lys Asp Asp Ser Gly Gln Arg Tyr Leu
    450                 455                 460

Ala Trp Ser Gly His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala
465                 470                 475                 480

Arg Ala Gly Val Ser Ile Pro Glu Ile Met Gln Ala Gly Gly Trp Thr
                485                 490                 495

Asn Val Asn Ile Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr
            500                 505                 510

Gly Ala Met Val Arg Leu Leu Glu Asp Gly Asp Gly Ile Glu Gly Arg
        515                 520                 525

Gly Ser Ala Trp Arg His Pro Gln Phe Gly Gly
    530                 535
```

<210> 15
<211> 5953
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: vector
pCRT7-deltaVPCS

<400> 15
cgatggattt ccgtctctgg tgtagctgat gatccgaata actacctgtt ttgccgggtc 60
agaaaaaatg gtgttgccgc gccatctgcc accagccagc tatcaactcg cgccctggaa 120
gggatttttg aagcaactca tcgattgatt tacggcgcta aggatgactc tggtcagaga 180
tacctggcct ggtctggaca cagtgcccgt gtcggagccg cgcgagatat ggcccgcgct 240
ggagtttcaa taccggagat catgcaagct ggtggctgga ccaatgtaaa tattgtcatg 300
aactatatcc gtaacctgga tagtgaaaca ggggcaatgg tgcgcctgct ggaagatggc 360
gatggtatcg aaggtcgtgg tagcgcttgg cgtcacccgc agttcggtgg ttaataagct 420
tcgaacaaaa actcatctca gaagaggatc tgaatatgca taccggtcat catcaccatc 480
accattgagt tttgagcaat aactagcata accccttggg gcctctaaac gggtcttgag 540
gggttttttg ctgaaaggag gaactatatc cggatatcca caggacgggt gtggtcgcca 600
tgatcgcgta gtcgatagtg gctccaagta gcgaagcgag caggactggg cggcggccaa 660
agcggtcgga cagtgctccg agaacgggtg cgcatagaaa ttgcatcaac gcatatagcg 720
ctagcagcac gccatagtga ctggcgatgc tgtcggaatg gacgatatcc cgcaagaggc 780
ccggcagtac cggcataacc aagcctatgc ctacagcatc cagggtgacg gtgccgagga 840
tgacgatgag cgcattgtta gatttcatac acggtgcctg actgcgttag caatttaact 900
gtgataaact accgcattaa agcttatcga tgataagctg tcaaacatga gaattaattc 960
ttagaaaaac tcatcgagca tcaaatgaaa ctgcaattta ttcatatcag gattatcaat 1020
accatatttt tgaaaaagcc gtttctgtaa tgaaggagaa aactcaccga ggcagttcca 1080
taggatggca agatcctggt atcggtctgc gattccgact cgtccaacat caatacaacc 1140
tattaatttc ccctcgtcaa aaataaggtt atcaagtgag aaatcaccat gagtgacgac 1200
tgaatccggt gagaatggca aaagcttatg catttctttc cagacttgtt caacaggcca 1260
gccattacgc tcgtcatcaa aatcactcgc atcaaccaaa ccgttattca ttcgtgattg 1320
cgcctgagcg agacgaaata cgcgatcgct gttaaaagga caattacaaa caggaatcga 1380
atgcaaccgg cgcaggaaca ctgccagcgc atcaacaata ttttcacctg aatcaggata 1440
ttcttctaat acctggaatg ctgttttccc ggggatcgca gtggtgagta accatgcatc 1500
atcaggagta cggataaaat gcttgatggt cggaagaggc ataaattccg tcagccagtt 1560
tagtctgacc atctcatctg taacatcatt ggcaacgcta cctttgccat gtttcagaaa 1620
caactctggc gcatcgggct tcccatacaa tcgatagatt gtcgcacctg attgcccgac 1680
attatcgcga gcccatttat acccatataa atcagcatcc atgttggaat ttaatcgcgg 1740
cctcgagcaa gacgtttccc gttgaatatg gctcataaca ccccttgtat tactgtttat 1800
gtaagcagac agttttattg ttcatgacca aaatccctta acgtgagttt tcgttccact 1860
gagcgtcaga ccccgtagaa aagatcaaag gatcttcttg agatcctttt tttctgcgcg 1920
taatctgctg cttgcaaaca aaaaaaccac cgctaccagc ggtggtttgt ttgccggatc 1980
aagagctacc aactcttttt ccgaaggtaa ctggcttcag cagagcgcag ataccaaata 2040
ctgtccttct agtgtagccg tagttaggcc accacttcaa gaactctgta gcaccgccta 2100
catacctcgc tctgctaatc ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc 2160
ttaccgggtt ggactcaaga cgatagttac cggataaggc gcagcggtcg ggctgaacgg 2220
ggggttcgtg cacacagccc agcttggagc gaacgaccta caccgaactg agatacctac 2280
agcgtgagct atgagaaagc gccacgcttc ccgaagggag aaaggcggac aggtatccgg 2340
taagcggcag ggtcggaaca ggagagcgca cgagggagct tccaggggga aacgcctggt 2400
atctttatag tcctgtcggg tttcgccacc tctgacttga gcgtcgattt ttgtgatgct 2460
cgtcaggggg gcggagccta tggaaaaacg ccagcaacgc ggccttttta cggttcctgg 2520
ccttttgctg gccttttgct cacatgttct ttcctgcgtt atcccctgat tctgtggata 2580
accgtattac cgcctttgag tgagctgata ccgctcgccg cagccgaacg accgagcgca 2640
gcgagtcagt gagcgaggaa gcggaagagc gcctgatgcg gtattttctc cttacgcatc 2700
tgtgcggtat ttcacaccgc atatatggtg cactctcagt acaatctgct ctgatgccgc 2760
atagttaagc cagtatacac tccgctatcg ctacgtgact gggtcatggc tgcgccccga 2820
cacccgccaa cacccgctga cgcgccctga cgggcttgtc tgctcccggc atccgcttac 2880
agacaagctg tgaccgtctc cgggagctgc atgtgtcaga ggttttcacc gtcatcaccg 2940
aaacgcgcga ggcagctgcg gtaaagctca tcagcgtggt cgtgaagcga ttcacagatg 3000
tctgcctgtt catccgcgtc cagctcgttg agtttctcca gaagcgttaa tgtctggctt 3060
ctgataaagc gggccatgtt aagggcggtt ttttcctgtt tggtcactga tgcctccgtg 3120
taagggggat ttctgttcat gggggtaatg ataccgatga aacgagagag gatgctcacg 3180
atacgggtta ctgatgatga acatgcccgg ttactggaac gttgtgaggg taaacaactg 3240

gcggtatgga tgcggcggga ccagagaaaa atcactcagg gtcaatgcca gcgcttcgtt 3300
aatacagatg taggtgttcc acagggtagc cagcagcatc ctgcgatgca gatccggaac 3360
ataatggtgc agggcgctga cttccgcgtt tccagacttt acgaaacacg gaaaccgaag 3420
accattcatg ttgttgctca ggtcgcagac gttttgcagc agcagtcgct tcacgttcgc 3480
tcgcgtatcg gtgattcatt ctgctaacca gtaaggcaac cccgccagcc tagccgggtc 3540
ctcaacgaca ggagcacgat catgcgcacc cgtggccagg acccaacgct gcccgagatg 3600
cgccgcgtgc ggctgctgga gatggcggac gcgatggata tgttctgcca agggttggtt 3660
tgcgcattca cagttctccg caagaattga ttggctccaa ttcttggagt ggtgaatccg 3720
ttagcgaggt gccgccggct tccattcagg tcgaggtggc ccggctccat gcaccgcgac 3780
gcaacgcggg gaggcagaca aggtataggg cggcgcctac aatccatgcc aacccgttcc 3840
atgtgctcgc cgaggcggca taaatcgccg tgacgatcag cggtccagtg atcgaagtta 3900
ggctggtaag agccgcgagc gatccttgaa gctgtccctg atggtcgtca tctacctgcc 3960
tggacagcat ggcctgcaac gcgggcatcc cgatgccgcc ggaagcgaga agaatcataa 4020
tggggaaggc catccagcct cgcgtcgcga acgccagcaa gacgtagccc agcgcgtcgg 4080
ccgccatgcc ggcgataatg gcctgcttct cgccgaaacg tttggtggcg ggaccagtga 4140
cgaaggcttg agcgagggcg tgcaagattc cgaataccgc aagcgacagg ccgatcatcg 4200
tcgcgctcca gcgaaagcgg tcctcgccga aaatgaccca gagcgctgcc ggcacctgtc 4260
ctacgagttg catgataaag aagacagtca taagtgcggc gacgatagtc atgccccgcg 4320
cccaccggaa ggagctgact gggttgaagg ctctcaaggg catcggtcga cgctctccct 4380
tatgcgactc ctgcattagg aagcagccca gtagtaggtt gaggccgttg agcaccgccg 4440
ccgcaaggaa tggtgcatgc aaggagatgg cgcccaacag tcccccggcc acggggcctg 4500
ccaccatacc cacgccgaaa caagcgctca tgagcccgaa gtggcgagcc cgatcttccc 4560
catcggtgat gtcggcgata taggcgccag caaccgcacc tgtggcgccg gtgatgccgg 4620
ccacgatgcg tccggcgtag aggatcgaga tctcgatccc gcgaaattaa tacgactcac 4680
tatagggaga ccacaacggt ttccctctag aaataatttt gtttaacttt aagaaggaga 4740
tatacatatg gctagcatga ctggtggaca gcaaatgggt cgggatccgt cgacggcgcc 4800
aaccccgatcc aagacacccg cgcaggggct ggccagaaag ctgcacttta gcaccgcccc 4860
cccaaacccc gacgcgccat ggaccccccg ggtggccggc tttaacaagc gcgtcttctg 4920
cgccgcggtc gggcgcctgg cggccatgca tgcccggatg gcggctgtcc agctctggga 4980
catgtcgcgt ccgcgcacag acgaagacct caacgaactc cttggcatca ccaccatccg 5040
cgtgacggtc tgcgagggca aaaacctgct tcagcgcgcc aacgagttgg tgaatccaga 5100
cgtggtgcag gacgtcgacg cggccacggc gactcgaggg cgttctgcgg cgtcgcgccc 5160
caccgagcga cctcgagccc cagcccgctc cgcttctcgc cccagacggc ccgtcgaggg 5220
taccgagctc ggatccacta gtccagtgtg gtggaattct gcagatatcc agcacagtgg 5280
cggccgcatg tccaatttac tgaccgtaca ccaaaatttg cctgcattac cggtcgatgc 5340
aacgagtgat gaggttcgca agaacctgat ggacatgttc agggatcgcc aggcgttttc 5400
tgagcatacc tggaaaatgc ttctgtccgt ttgccggtcg tgggcggcat ggtgcaagtt 5460
gaataaccgg aaatggtttc ccgcagaacc tgaagatgtt cgcgattatc ttctatatct 5520
tcaggcgcgc ggtctggcag taaaaactat ccagcaacat ttgggccagc taaacatgct 5580
tcatcgtcgg tccgggctgc cacgaccaag tgacagcaat gctgtttcac tggttatgcg 5640
gcggatccga aaagaaaacg ttgatgccgg tgaacgtgca aaacaggctc tagcgttcga 5700
acgcactgat ttcgaccagg ttcgttcact catggaaaat agcgatcgct gccaggatat 5760
acgtaatctg gcatttctgg ggattgctta taacaccctg ttacgtatag ccgaaattgc 5820
caggatcagg gttaaagata tctcacgtac tgacggtggg agaatgttaa tccatattgg 5880
cagaacgaaa acgctggtta gcaccgcagg tgtagagaag gcacttagcc tgggggtaac 5940
taaactggtc gag 5953

<210> 16
<211> 4727
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: vector pT7-TACS

<400> 16
atccggatat agttcctcct ttcagcaaaa aacccctcaa gacccgttta gaggccccaa 60
ggggttatgc tagttattgc tcagcggtgg cagcagccaa ctcagcttcc tttcgggctt 120
tgttagcagc cggatctcag tggtggtggt ggtggtgctc gagtgcggcc gcaagcttat 180
taaccaccga actgcgggtg acgccaagcg ctaccacgac cttcgatacc atcgccatct 240
tccagcaggc gcaccattgc ccctgtttca ctatccaggt tacggatata gttcatgaca 300
atatttacat tggtccagcc accagcttgc atgatctccg gtattgaaac tccagcgcgg 360

```
gccatatctc gcgcggctcc gacacgggca ctgtgtccag accaggccag gtatctctga 420
ccagagtcat ccttagcgcc gtaaatcaat cgatgagttg cttcaaaaat cccttccagg 480
gcgcgagttg atagctggct ggtggcagat ggcgcggcaa caccattttt tctgacccgg 540
caaaacaggt agttattcgg atcatcagct acaccagaga cggaaatcca tcgctcgacc 600
agtttagtta cccccaggct aagtgccttc tctacacctg cggtgctaac cagcgttttc 660
gttctgccaa tatggattaa cattctccca ccgtcagtac gtgagatatc tttaaccctg 720
atcctggcaa tttcggctat acgtaacagg gtgttataag caatccccag aaatgccaga 780
ttacgtatat cctggcagcg atcgctattt tccatgagtg aacgaacctg gtcgaaatca 840
gtgcgttcga acgctagagc ctgttttgca cgttcaccgg catcaacgtt ttcttttcgg 900
atccgccgca taaccagtga aacagcattg ctgtcacttg gtcgtggcag cccggaccga 960
cgatgaagca tgtttagctg gcccaaatgt tgctggatag tttttactgt cagaccgcgc 1020
gcctgaagat atagaagata atcgcgaaca tcttcaggtt ctgcgggaaa ccatttccgg 1080
ttattcaact tgcaccatgc cgcccacgac cggcaaacgg acagaagcat tttccaggta 1140
tgctcagaaa acgcctggcg atccctgaac atgtccatca ggttcttgcg aacctcatca 1200
ctcgttgcat cgaccggtaa tgcaggcaaa ttttggtgta cggtcagtaa attggacatg 1260
ccgcggcggc gttggcggcg cttcttgcgg ccgtagccca tggtatatct ccttcttaaa 1320
gttaaacaaa attatttcta gagggaaacc gttgtggtct ccctatagtg agtcgtatta 1380
atttcgcggg atcgagatct cgggcagcgt tgggtcctgg ccacgggtgc gcatgatcgt 1440
gctcctgtcg ttgaggaccc ggctaggctg gcggggttgc cttactggtt agcagaatga 1500
atcaccgata cgcgagcgaa cgtgaagcga ctgctgctgc aaaacgtctg cgacctgagc 1560
aacaacatga atggtcttcg gtttccgtgt ttcgtaaagt ctggaaacgc ggaagtcagc 1620
gccctgcacc attatgttcc ggatctgcat cgcaggatgc tgctggctac cctgtggaac 1680
acctacatct gtattaacga agcgctggca ttgaccctga gtgatttttc tctggtcccg 1740
ccgcatccat accgccagtt gtttaccctc acaacgttcc agtaaccggg catgttcatc 1800
atcagtaacc cgtatcgtga gcatcctctc tcgtttcatc ggtatcatta cccccatgaa 1860
cagaaatccc ccttacacgg aggcatcagt gaccaaacag gaaaaaaccg cccttaacat 1920
ggcccgcttt atcagaagcc agacattaac gcttctggag aaactcaacg agctggacgc 1980
ggatgaacag gcagacatct gtgaatcgct tcacgaccac gctgatgagc tttaccgcag 2040
ctgcctcgcg cgtttcggtg atgacggtga aaacctctga cacatgcagc tcccggagac 2100
ggtcacagct tgtctgtaag cggatgccgg gagcagacaa gcccgtcagg gcgcgtcagc 2160
gggtgttggc gggtgtcggg gcgcagccat gacccagtca cgtagcgata gcggagtgta 2220
tactggctta actatgcggc atcagagcag attgtactga gagtgcacca tatatgcggt 2280
gtgaaatacc gcacagatgc gtaaggagaa aataccgcat caggcgctct tccgcttcct 2340
cgctcactga ctcgctgcgc tcggtcgttc ggctgcggcg agcggtatca gctcactcaa 2400
aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac atgtgagcaa 2460
aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc 2520
tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga 2580
caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc 2640
cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc gtggcgcttt 2700
ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc aagctgggct 2760
gtgtgcacga accccccgtt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg 2820
agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta 2880
gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct 2940
acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa 3000
gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt ttttttgttt 3060
gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta 3120
cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc atgagattat 3180
caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa tcaatctaaa 3240
gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag gcacctatct 3300
cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg tagataacta 3360
cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga gacccacgct 3420
caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag cgcagaagtg 3480
gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa gctagagtaa 3540
gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctgcaggc atcgtggtgt 3600
cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca aggcgagtta 3660
catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg atcgttgtca 3720
gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat aattctctta 3780
ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc aagtcattct 3840
gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg gataataccg 3900
cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg gggcgaaaac 3960
tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt gcacccaact 4020
gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca ggaaggcaaa 4080
atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata ctcttccttt 4140
ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac atatttgaat 4200
```

gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa gtgccacctg 4260
aaattgtaaa cgttaatatt ttgttaaaat tcgcgttaaa tttttgttaa atcagctcat 4320
tttttaacca ataggccgaa atcggcaaaa tcccttataa atcaaaagaa tagaccgaga 4380
tagggttgag tgttgttcca gtttggaaca agagtccact attaaagaac gtggactcca 4440
acgtcaaagg gcgaaaaacc gtctatcagg gcgatggccc actacgtgaa ccatcaccct 4500
aatcaagttt tttggggtcg aggtgccgta aagcactaaa tcggaaccct aaagggagcc 4560
cccgatttag agcttgacgg ggaaagccgg cgaacgtggc gagaaaggaa gggaagaaag 4620
cgaaaggagc gggcgctagg gcgctggcaa gtgtagcggt cacgctgcgc gtaaccacca 4680
cacccgccgc gcttaatgcg ccgctacagg gcgcgtccca ttcgcca 4727

<210> 17
<211> 4488
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: vector
pT7-VPCS

<400> 17
aaatcaatct aaagtatata tgagtaaact tggtctgaca gttaccaatg cttaatcagt 60
gaggcaccta tctcagcgat ctgtctattt cgttcatcca tagttgcctg actccccgtc 120
gtgtagataa ctacgatacg ggagggctta ccatctggcc ccagtgctgc aatgataccg 180
cgagacccac gctcaccggc tccagattta tcagcaataa accagccagc cggaagggcc 240
gagcgcagaa gtggtcctgc aactttatcc gcctccatcc agtctattaa ttgttgccgg 300
gaagctagag taagtagttc gccagttaat agtttgcgca acgttgttgc cattgctaca 360
ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat tcagctccgg ttcccaacga 420
tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag cggttagctc cttcggtcct 480
ccgatcgttg tcagaagtaa gttggccgca gtgttatcac tcatggttat ggcagcactg 540
cataattctc ttactgtcat gccatccgta agatgctttt ctgtgactgg tgagtactca 600
accaagtcat tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaaca 660
cgggataata ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg aaaacgttct 720
tcggggcgaa aactctcaag gatcttaccg ctgttgagat ccagttcgat gtaacccact 780
cgtgcaccca actgatcttc agcatctttt actttcacca gcgtttctgg gtgagcaaaa 840
acaggaaggc aaaatgccgc aaaaaaggga ataagggcga cacggaaatg ttgaatactc 900
atactcttcc tttttcaata ttattgaagc atttatcagg gttattgtct catgagcgga 960
tacatatttg aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga 1020
aaagtgccac ctgacgtcta agaaaccatt attatcatga cattaaccta taaaaatagg 1080
cgtatcacga ggccctttcg tcttcaagaa ttaaaaggat ctaggtgaag atcctttttg 1140
ataatctcat gaccaaaatc ccttaacgtg agttttcgtt ccactgagcg tcagaccccg 1200
tagaaaagat caaaggatct tcttgagatc cttttttct gcgcgtaatc tgctgcttgc 1260
aaacaaaaaa accaccgcta ccagcggtgg tttgtttgcc ggatcaagag ctaccaactc 1320
tttttccgaa ggtaactggc ttcagcagag cgcagatacc aaatactgtc cttctagtgt 1380
agccgtagtt aggccaccac ttcaagaact ctgtagcacc gcctacatac ctcgctctgc 1440
taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc gggttggact 1500
caagacgata gttaccggat aaggcgcagc ggtcgggctg aacggggggt tcgtgcacac 1560
agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag 1620
aaagcgccac gcttcccgaa gggagaaagg cggacaggta tccggtaagc ggcagggtcg 1680
gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggtatctt tatagtcctg 1740
tcgggtttcg ccacctctga cttgagcgtc gatttttgtg atgctcgtca ggggggcgga 1800
gcctatggaa aaacgccagc aacgcggcct ttttacggtt cctggccttt tgctggcctt 1860
ttgctcacat gttctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct 1920
ttgagtgagc tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg 1980
aggaagcgga agagcgcctg atgcggtatt ttctccttac gcatctgtgc ggtatttcac 2040
accgcatcag atctgatggt gcactctcag tacaatctgc tctgatgccg catagttaag 2100
ccagtatata cactccgcta tcgctacgtg actgggtcat ggctgcgccc cgacacccgc 2160
caacacccgc tgacgcgccc tgacgggctt gtctgctccc ggcatccgct tacagacaag 2220
ctgtgaccgt ctccgggagc tgcatgtgtc agaggttttc accgtcatca ccgaaacgcg 2280
cgaggcccag cgattcgaac ttctgataga cttcgaaatt aatacgactc actataggga 2340
gaccacaacg gtttccctct agaaataatt ttgtttaact ttaagaagga gatatacata 2400
tgacctctcg ccgctccgtg aagtcgggtc cgcgggaggt tccgcgcgat gagtacgagg 2460
atctgtacta caccccgtct tcaggtatgg cgagtcccga tagtccgcct gacacctccc 2520
gccgtggcgc cctacagaca cgctcgcgcc agaggggcga ggtccgtttc gtccagtacg 2580

acgagtcgga ttatgccctc tacgggggct cgtcttccga agacgacgaa cacccggagg 2640
tccccccggac gcggcgtccc gtttccgggg cggttttgtc cggcccgggg cctgcgcggg 2700
cgcctccgcc acccgctggg tccggagggg ccggacgcac acccaccacc gccccccggg 2760
ccccccgaac ccagcgggtg gcgtctaagg cccccgcggc cccggcggcg gagaccaccc 2820
gcggcaggaa atcggcccag ccagaatccg ccgcactccc agacgccccc gcgtcgacgg 2880
cgccaacccg atccaagaca cccgcgcagg ggctggccag aaagctgcac tttagcaccg 2940
cccccccaaa ccccgacgcg ccatggaccc cccgggtggc cggctttaac aagcgcgtct 3000
tctgcgccgc ggtcgggcgc ctggcggcca tgcatgcccg gatggcggct gtccagctct 3060
gggacatgtc gcgtccgcgc acagacgaag acctcaacga actccttggc atcaccacca 3120
tccgcgtgac ggtctgcgag ggcaaaaacc tgcttcagcg cgccaacgag ttggtgaatc 3180
cagacgtggt gcaggacgtc gacgcggcca cggcgactcg agggcgttct gcggcgtcgc 3240
gccccaccga gcgacctcga gccccagccc gctccgcttc tcgccccaga cggcccgtcg 3300
agggtaccga gctcggatcc actagtccag tgtggtggaa ttctgcagat atccagcaca 3360
gtggcggccg catgtccaat ttactgaccg tacaccaaaa tttgcctgca ttaccggtcg 3420
atgcaacgag tgatgaggtt cgcaagaacc tgatggacat gttcagggat cgccaggcgt 3480
tttctgagca tacctggaaa atgcttctgt ccgtttgccg gtcgtgggcg gcatggtgca 3540
agttgaataa ccggaaatgg tttcccgcag aacctgaaga tgttcgcgat tatcttctat 3600
atcttcaggc gcgcggtctg gcagtaaaaa ctatccagca acatttgggc cagctaaaca 3660
tgcttcatcg tcggtccggg ctgccacgac caagtgacag caatgctgtt tcactggtta 3720
tgcggcggat ccgaaaagaa aacgttgatg ccggtgaacg tgcaaaacag gctctagcgt 3780
tcgaacgcac tgatttcgac caggttcgtt cactcatgga aaatagcgat cgctgccagg 3840
atatacgtaa tctggcattt ctggggattg cttataacac cctgttacgt atagccgaaa 3900
ttgccaggat cagggttaaa gatatctcac gtactgacgg tgggagaatg ttaatccata 3960
ttggcagaac gaaaacgctg gttagcaccg caggtgtaga gaaggcactt agcctggggg 4020
taactaaact ggtcgagcga tggatttccg tctctggtgt agctgatgat ccgaataact 4080
acctgttttg ccgggtcaga aaaaatggtg ttgccgcgcc atctgccacc agccagctat 4140
caactcgcgc cctggaaggg atttttgaag caactcatcg attgatttac ggcgctaagg 4200
atgactctgg tcagagatac ctggcctggt ctggacacag tgcccgtgtc ggagccgcgc 4260
gagatatggc ccgcgctgga gtttcaatac cggagatcat gcaagctggt ggctggacca 4320
atgtaaatat tgtcatgaac tatatccgta acctggatag tgaaacaggg gcaatggtgc 4380
gcctgctgga agatggcgat ggtatcgaag gtcgtggtag cgcttggcgt cacccgcagt 4440
tcggtggtta ataagcttat cgatgataag ctgtcaaaca tgagaatt 4488

<210> 18
<211> 1125
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA sequence
coding for a fusion protein TATcreStrepTag

<220>
<221> CDS
<222> (1)..(1119)

<400> 18
atg ggc tac ggc cgc aag aag cgc cgc caa cgc cgc cgc ggc atg tcc 48
Met Gly Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Gly Met Ser
1 5 10 15

aat tta ctg acc gta cac caa aat ttg cct gca tta ccg gtc gat gca 96
Asn Leu Leu Thr Val His Gln Asn Leu Pro Ala Leu Pro Val Asp Ala
20 25 30

acg agt gat gag gtt cgc aag aac ctg atg gac atg ttc agg gat cgc 144
Thr Ser Asp Glu Val Arg Lys Asn Leu Met Asp Met Phe Arg Asp Arg
35 40 45

cag gcg ttt tct gag cat acc tgg aaa atg ctt ctg tcc gtt tgc cgg 192
Gln Ala Phe Ser Glu His Thr Trp Lys Met Leu Leu Ser Val Cys Arg
50 55 60

```
tcg tgg gcg gca tgg tgc aag ttg aat aac cgg aaa tgg ttt ccc gca    240
Ser Trp Ala Ala Trp Cys Lys Leu Asn Asn Arg Lys Trp Phe Pro Ala
 65                 70                  75                  80

gaa cct gaa gat gtt cgc gat tat ctt cta tat ctt cag gcg cgc ggt    288
Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu Tyr Leu Gln Ala Arg Gly
                 85                  90                  95

ctg aca gta aaa act atc cag caa cat ttg ggc cag cta aac atg ctt    336
Leu Thr Val Lys Thr Ile Gln Gln His Leu Gly Gln Leu Asn Met Leu
            100                 105                 110

cat cgt cgg tcc ggg ctg cca cga cca agt gac agc aat gct gtt tca    384
His Arg Arg Ser Gly Leu Pro Arg Pro Ser Asp Ser Asn Ala Val Ser
        115                 120                 125

ctg gtt atg cgg cgg atc cga aaa gaa aac gtt gat gcc ggt gaa cgt    432
Leu Val Met Arg Arg Ile Arg Lys Glu Asn Val Asp Ala Gly Glu Arg
    130                 135                 140

gca aaa cag gct cta gcg ttc gaa cgc act gat ttc gac cag gtt cgt    480
Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Phe Asp Gln Val Arg
145                 150                 155                 160

tca ctc atg gaa aat agc gat cgc tgc cag gat ata cgt aat ctg gca    528
Ser Leu Met Glu Asn Ser Asp Arg Cys Gln Asp Ile Arg Asn Leu Ala
                165                 170                 175

ttt ctg ggg att gct tat aac acc ctg tta cgt ata gcc gaa att gcc    576
Phe Leu Gly Ile Ala Tyr Asn Thr Leu Leu Arg Ile Ala Glu Ile Ala
            180                 185                 190

agg atc agg gtt aaa gat atc tca cgt act gac ggt ggg aga atg tta    624
Arg Ile Arg Val Lys Asp Ile Ser Arg Thr Asp Gly Gly Arg Met Leu
        195                 200                 205

atc cat att ggc aga acg aaa acg ctg gtt agc acc gca ggt gta gag    672
Ile His Ile Gly Arg Thr Lys Thr Leu Val Ser Thr Ala Gly Val Glu
    210                 215                 220

aag gca ctt agc ctg ggg gta act aaa ctg gtc gag cga tgg att tcc    720
Lys Ala Leu Ser Leu Gly Val Thr Lys Leu Val Glu Arg Trp Ile Ser
225                 230                 235                 240

gtc tct ggt gta gct gat gat ccg aat aac tac ctg ttt tgc cgg gtc    768
Val Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys Arg Val
                245                 250                 255

aga aaa aat ggt gtt gcc gcg cca tct gcc acc agc cag cta tca act    816
Arg Lys Asn Gly Val Ala Ala Pro Ser Ala Thr Ser Gln Leu Ser Thr
            260                 265                 270

cgc gcc ctg gaa ggg att ttt gaa gca act cat cga ttg att tac ggc    864
Arg Ala Leu Glu Gly Ile Phe Glu Ala Thr His Arg Leu Ile Tyr Gly
        275                 280                 285

gct aag gat gac tct ggt cag aga tac ctg gcc tgg tct gga cac agt    912
Ala Lys Asp Asp Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly His Ser
    290                 295                 300

gcc cgt gtc gga gcc gcg cga gat atg gcc cgc gct gga gtt tca ata    960
Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val Ser Ile
305                 310                 315                 320
```

ccg gag atc atg caa gct ggt ggc tgg acc aat gta aat att gtc atg 1008
Pro Glu Ile Met Gln Ala Gly Gly Trp Thr Asn Val Asn Ile Val Met
325 330 335

aac tat atc cgt aac ctg gat agt gaa aca ggg gca atg gtg cgc ctg 1056
Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val Arg Leu
340 345 350

ctg gaa gat ggc gat ggt atc gaa ggt cgt ggt agc gct tgg cgt cac 1104
Leu Glu Asp Gly Asp Gly Ile Glu Gly Arg Gly Ser Ala Trp Arg His
355 360 365

ccg cag ttc ggt ggt taataa 1125
Pro Gln Phe Gly Gly
370

<210> 19
<211> 373
<212> PRT
<213> Artificial Sequence
<223> Description of Artificial Sequence: DNA sequence coding for a fusion protein TATcreStrepTag

<400> 19
Met Gly Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Gly Met Ser
1 5 10 15

Asn Leu Leu Thr Val His Gln Asn Leu Pro Ala Leu Pro Val Asp Ala
20 25 30

Thr Ser Asp Glu Val Arg Lys Asn Leu Met Asp Met Phe Arg Asp Arg
35 40 45

Gln Ala Phe Ser Glu His Thr Trp Lys Met Leu Leu Ser Val Cys Arg
50 55 60

Ser Trp Ala Ala Trp Cys Lys Leu Asn Asn Arg Lys Trp Phe Pro Ala
65 70 75 80

Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu Tyr Leu Gln Ala Arg Gly
85 90 95

Leu Thr Val Lys Thr Ile Gln Gln His Leu Gly Gln Leu Asn Met Leu
100 105 110

His Arg Arg Ser Gly Leu Pro Arg Pro Ser Asp Ser Asn Ala Val Ser
115 120 125

Leu Val Met Arg Arg Ile Arg Lys Glu Asn Val Asp Ala Gly Glu Arg
130 135 140

Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Phe Asp Gln Val Arg
145 150 155 160

Ser Leu Met Glu Asn Ser Asp Arg Cys Gln Asp Ile Arg Asn Leu Ala
165 170 175

Phe Leu Gly Ile Ala Tyr Asn Thr Leu Leu Arg Ile Ala Glu Ile Ala
180 185 190

Arg Ile Arg Val Lys Asp Ile Ser Arg Thr Asp Gly Gly Arg Met Leu
195 200 205

Ile His Ile Gly Arg Thr Lys Thr Leu Val Ser Thr Ala Gly Val Glu
    210   215   220

Lys Ala Leu Ser Leu Gly Val Thr Lys Leu Val Glu Arg Trp Ile Ser
225   230   235   240

Val Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys Arg Val
                245   250   255

Arg Lys Asn Gly Val Ala Ala Pro Ser Ala Thr Ser Gln Leu Ser Thr
            260   265   270

Arg Ala Leu Glu Gly Ile Phe Glu Ala Thr His Arg Leu Ile Tyr Gly
        275   280   285

Ala Lys Asp Asp Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly His Ser
    290   295   300

Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val Ser Ile
305   310   315   320

Pro Glu Ile Met Gln Ala Gly Gly Trp Thr Asn Val Asn Ile Val Met
                325   330   335

Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val Arg Leu
            340   345   350

Leu Glu Asp Gly Asp Gly Ile Glu Gly Arg Gly Ser Ala Trp Arg His
        355   360   365

Pro Gln Phe Gly Gly
    370

<210> 20
<211> 2055
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA sequence
coding for a fusion protein VP22creStrepTag

<220>
<221> CDS
<222> (1)..(2049)

<400> 20
atg acc tct cgc cgc tcc gtg aag tcg ggt ccg cgg gag gtt ccg cgc 48
Met Thr Ser Arg Arg Ser Val Lys Ser Gly Pro Arg Glu Val Pro Arg
1   5   10   15

gat gag tac gag gat ctg tac tac acc ccg tct tca ggt atg gcg agt 96
Asp Glu Tyr Glu Asp Leu Tyr Tyr Thr Pro Ser Ser Gly Met Ala Ser
            20   25   30

ccc gat agt ccg cct gac acc tcc cgc cgt ggc gcc cta cag aca cgc 144
Pro Asp Ser Pro Pro Asp Thr Ser Arg Arg Gly Ala Leu Gln Thr Arg
        35   40   45

tcg cgc cag agg ggc gag gtc cgt ttc gtc cag tac gac gag tcg gat 192
Ser Arg Gln Arg Gly Glu Val Arg Phe Val Gln Tyr Asp Glu Ser Asp
    50   55   60

```
tat gcc ctc tac ggg ggc tcg tct tcc gaa gac gac gaa cac ccg gag   240
Tyr Ala Leu Tyr Gly Gly Ser Ser Ser Glu Asp Asp Glu His Pro Glu
 65                  70                  75                  80

gtc ccc cgg acg cgg cgt ccc gtt tcc ggg gcg gtt ttg tcc ggc ccg   288
Val Pro Arg Thr Arg Arg Pro Val Ser Gly Ala Val Leu Ser Gly Pro
                 85                  90                  95

ggg cct gcg cgg gcg cct ccg cca ccc gct ggg tcc gga ggg gcc gga   336
Gly Pro Ala Arg Ala Pro Pro Pro Pro Ala Gly Ser Gly Gly Ala Gly
            100                 105                 110

cgc aca ccc acc acc gcc ccc cgg gcc ccc cga acc cag cgg gtg gcg   384
Arg Thr Pro Thr Thr Ala Pro Arg Ala Pro Arg Thr Gln Arg Val Ala
        115                 120                 125

tct aag gcc ccc gcg gcc ccg gcg gcg gag acc acc cgc ggc agg aaa   432
Ser Lys Ala Pro Ala Ala Pro Ala Ala Glu Thr Thr Arg Gly Arg Lys
    130                 135                 140

tcg gcc cag cca gaa tcc gcc gca ctc cca gac gcc ccc gcg tcg acg   480
Ser Ala Gln Pro Glu Ser Ala Ala Leu Pro Asp Ala Pro Ala Ser Thr
145                 150                 155                 160

gcg cca acc cga tcc aag aca ccc gcg cag ggg ctg gcc aga aag ctg   528
Ala Pro Thr Arg Ser Lys Thr Pro Ala Gln Gly Leu Ala Arg Lys Leu
                165                 170                 175

cac ttt agc acc gcc ccc cca aac ccc gac gcg cca tgg acc ccc cgg   576
His Phe Ser Thr Ala Pro Pro Asn Pro Asp Ala Pro Trp Thr Pro Arg
            180                 185                 190

gtg gcc ggc ttt aac aag cgc gtc ttc tgc gcc gcg gtc ggg cgc ctg   624
Val Ala Gly Phe Asn Lys Arg Val Phe Cys Ala Ala Val Gly Arg Leu
        195                 200                 205

gcg gcc atg cat gcc cgg atg gcg gct gtc cag ctc tgg gac atg tcg   672
Ala Ala Met His Ala Arg Met Ala Ala Val Gln Leu Trp Asp Met Ser
    210                 215                 220

cgt ccg cgc aca gac gaa gac ctc aac gaa ctc ctt ggc atc acc acc   720
Arg Pro Arg Thr Asp Glu Asp Leu Asn Glu Leu Leu Gly Ile Thr Thr
225                 230                 235                 240

atc cgc gtg acg gtc tgc gag ggc aaa aac ctg ctt cag cgc gcc aac   768
Ile Arg Val Thr Val Cys Glu Gly Lys Asn Leu Leu Gln Arg Ala Asn
                245                 250                 255

gag ttg gtg aat cca gac gtg gtg cag gac gtc gac gcg gcc acg gcg   816
Glu Leu Val Asn Pro Asp Val Val Gln Asp Val Asp Ala Ala Thr Ala
            260                 265                 270

act cga ggg cgt tct gcg gcg tcg cgc ccc acc gag cga cct cga gcc   864
Thr Arg Gly Arg Ser Ala Ala Ser Arg Pro Thr Glu Arg Pro Arg Ala
        275                 280                 285

cca gcc cgc tcc gct tct cgc ccc aga cgg ccc gtc gag ggt acc gag   912
Pro Ala Arg Ser Ala Ser Arg Pro Arg Arg Pro Val Glu Gly Thr Glu
    290                 295                 300

ctc gga tcc act agt cca gtg tgg tgg aat tct gca gat atc cag cac   960
Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Gln His
305                 310                 315                 320
```

```
agt ggc ggc cgc atg tcc aat tta ctg acc gta cac caa aat ttg cct    1008
Ser Gly Gly Arg Met Ser Asn Leu Leu Thr Val His Gln Asn Leu Pro
                325                 330                 335

gca tta ccg gtc gat gca acg agt gat gag gtt cgc aag aac ctg atg    1056
Ala Leu Pro Val Asp Ala Thr Ser Asp Glu Val Arg Lys Asn Leu Met
            340                 345                 350

gac atg ttc agg gat cgc cag gcg ttt tct gag cat acc tgg aaa atg    1104
Asp Met Phe Arg Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Met
        355                 360                 365

ctt ctg tcc gtt tgc cgg tcg tgg gcg gca tgg tgc aag ttg aat aac    1152
Leu Leu Ser Val Cys Arg Ser Trp Ala Ala Trp Cys Lys Leu Asn Asn
    370                 375                 380

cgg aaa tgg ttt ccc gca gaa cct gaa gat gtt cgc gat tat ctt cta    1200
Arg Lys Trp Phe Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu
385                 390                 395                 400

tat ctt cag gcg cgc ggt ctg gca gta aaa act atc cag caa cat ttg    1248
Tyr Leu Gln Ala Arg Gly Leu Ala Val Lys Thr Ile Gln Gln His Leu
                405                 410                 415

ggc cag cta aac atg ctt cat cgt cgg tcc ggg ctg cca cga cca agt    1296
Gly Gln Leu Asn Met Leu His Arg Arg Ser Gly Leu Pro Arg Pro Ser
            420                 425                 430

gac agc aat gct gtt tca ctg gtt atg cgg cgg atc cga aaa gaa aac    1344
Asp Ser Asn Ala Val Ser Leu Val Met Arg Arg Ile Arg Lys Glu Asn
        435                 440                 445

gtt gat gcc ggt gaa cgt gca aaa cag gct cta gcg ttc gaa cgc act    1392
Val Asp Ala Gly Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr
    450                 455                 460

gat ttc gac cag gtt cgt tca ctc atg gaa aat agc gat cgc tgc cag    1440
Asp Phe Asp Gln Val Arg Ser Leu Met Glu Asn Ser Asp Arg Cys Gln
465                 470                 475                 480

gat ata cgt aat ctg gca ttt ctg ggg att gct tat aac acc ctg tta    1488
Asp Ile Arg Asn Leu Ala Phe Leu Gly Ile Ala Tyr Asn Thr Leu Leu
                485                 490                 495

cgt ata gcc gaa att gcc agg atc agg gtt aaa gat atc tca cgt act    1536
Arg Ile Ala Glu Ile Ala Arg Ile Arg Val Lys Asp Ile Ser Arg Thr
            500                 505                 510

gac ggt ggg aga atg tta atc cat att ggc aga acg aaa acg ctg gtt    1584
Asp Gly Gly Arg Met Leu Ile His Ile Gly Arg Thr Lys Thr Leu Val
        515                 520                 525

agc acc gca ggt gta gag aag gca ctt agc ctg ggg gta act aaa ctg    1632
Ser Thr Ala Gly Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu
    530                 535                 540

gtc gag cga tgg att tcc gtc tct ggt gta gct gat gat ccg aat aac    1680
Val Glu Arg Trp Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn
545                 550                 555                 560

tac ctg ttt tgc cgg gtc aga aaa aat ggt gtt gcc gcg cca tct gcc    1728
Tyr Leu Phe Cys Arg Val Arg Lys Asn Gly Val Ala Ala Pro Ser Ala
                565                 570                 575
```

acc agc cag cta tca act cgc gcc ctg gaa ggg att ttt gaa gca act 1776
Thr Ser Gln Leu Ser Thr Arg Ala Leu Glu Gly Ile Phe Glu Ala Thr
580 585 590

cat cga ttg att tac ggc gct aag gat gac tct ggt cag aga tac ctg 1824
His Arg Leu Ile Tyr Gly Ala Lys Asp Asp Ser Gly Gln Arg Tyr Leu
595 600 605

gcc tgg tct gga cac agt gcc cgt gtc gga gcc gcg cga gat atg gcc 1872
Ala Trp Ser Gly His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala
610 615 620

cgc gct gga gtt tca ata ccg gag atc atg caa gct ggt ggc tgg acc 1920
Arg Ala Gly Val Ser Ile Pro Glu Ile Met Gln Ala Gly Gly Trp Thr
625 630 635 640

aat gta aat att gtc atg aac tat atc cgt aac ctg gat agt gaa aca 1968
Asn Val Asn Ile Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr
645 650 655

ggg gca atg gtg cgc ctg ctg gaa gat ggc gat ggt atc gaa ggt cgt 2016
Gly Ala Met Val Arg Leu Leu Glu Asp Gly Asp Gly Ile Glu Gly Arg
660 665 670

ggt agc gct tgg cgt cac ccg cag ttc ggt ggt taataa 2055
Gly Ser Ala Trp Arg His Pro Gln Phe Gly Gly
675 680

<210> 21
<211> 683
<212> PRT
<213> Artificial Sequence
<223> Description of Artificial Sequence: DNA sequence
coding for a fusion protein VP22creStrepTag

<400> 21
Met Thr Ser Arg Arg Ser Val Lys Ser Gly Pro Arg Glu Val Pro Arg
1 5 10 15

Asp Glu Tyr Glu Asp Leu Tyr Tyr Thr Pro Ser Ser Gly Met Ala Ser
20 25 30

Pro Asp Ser Pro Pro Asp Thr Ser Arg Arg Gly Ala Leu Gln Thr Arg
35 40 45

Ser Arg Gln Arg Gly Glu Val Arg Phe Val Gln Tyr Asp Glu Ser Asp
50 55 60

Tyr Ala Leu Tyr Gly Gly Ser Ser Ser Glu Asp Asp Glu His Pro Glu
65 70 75 80

Val Pro Arg Thr Arg Arg Pro Val Ser Gly Ala Val Leu Ser Gly Pro
85 90 95

Gly Pro Ala Arg Ala Pro Pro Pro Pro Ala Gly Ser Gly Gly Ala Gly
100 105 110

Arg Thr Pro Thr Thr Ala Pro Arg Ala Pro Arg Thr Gln Arg Val Ala
115 120 125

Ser Lys Ala Pro Ala Ala Pro Ala Ala Glu Thr Thr Arg Gly Arg Lys
130 135 140

```
Ser Ala Gln Pro Glu Ser Ala Ala Leu Pro Asp Ala Pro Ala Ser Thr
145                 150                 155                 160

Ala Pro Thr Arg Ser Lys Thr Pro Ala Gln Gly Leu Ala Arg Lys Leu
                165                 170                 175

His Phe Ser Thr Ala Pro Pro Asn Pro Asp Ala Pro Trp Thr Pro Arg
            180                 185                 190

Val Ala Gly Phe Asn Lys Arg Val Phe Cys Ala Ala Val Gly Arg Leu
        195                 200                 205

Ala Ala Met His Ala Arg Met Ala Ala Val Gln Leu Trp Asp Met Ser
    210                 215                 220

Arg Pro Arg Thr Asp Glu Asp Leu Asn Glu Leu Leu Gly Ile Thr Thr
225                 230                 235                 240

Ile Arg Val Thr Val Cys Glu Gly Lys Asn Leu Leu Gln Arg Ala Asn
                245                 250                 255

Glu Leu Val Asn Pro Asp Val Val Gln Asp Val Asp Ala Ala Thr Ala
            260                 265                 270

Thr Arg Gly Arg Ser Ala Ala Ser Arg Pro Thr Glu Arg Pro Arg Ala
        275                 280                 285

Pro Ala Arg Ser Ala Ser Arg Pro Arg Arg Pro Val Glu Gly Thr Glu
    290                 295                 300

Leu Gly Ser Thr Ser Pro Val Trp Trp Asn Ser Ala Asp Ile Gln His
305                 310                 315                 320

Ser Gly Gly Arg Met Ser Asn Leu Leu Thr Val His Gln Asn Leu Pro
                325                 330                 335

Ala Leu Pro Val Asp Ala Thr Ser Asp Glu Val Arg Lys Asn Leu Met
            340                 345                 350

Asp Met Phe Arg Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Met
        355                 360                 365

Leu Leu Ser Val Cys Arg Ser Trp Ala Ala Trp Cys Lys Leu Asn Asn
    370                 375                 380

Arg Lys Trp Phe Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu
385                 390                 395                 400

Tyr Leu Gln Ala Arg Gly Leu Ala Val Lys Thr Ile Gln Gln His Leu
                405                 410                 415

Gly Gln Leu Asn Met Leu His Arg Arg Ser Gly Leu Pro Arg Pro Ser
            420                 425                 430

Asp Ser Asn Ala Val Ser Leu Val Met Arg Arg Ile Arg Lys Glu Asn
        435                 440                 445

Val Asp Ala Gly Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr
    450                 455                 460

Asp Phe Asp Gln Val Arg Ser Leu Met Glu Asn Ser Asp Arg Cys Gln
465                 470                 475                 480
```

Asp Ile Arg Asn Leu Ala Phe Leu Gly Ile Ala Tyr Asn Thr Leu Leu
                485   490   495

Arg Ile Ala Glu Ile Ala Arg Ile Arg Val Lys Asp Ile Ser Arg Thr
            500   505   510

Asp Gly Gly Arg Met Leu Ile His Ile Gly Arg Thr Lys Thr Leu Val
        515   520   525

Ser Thr Ala Gly Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu
    530   535   540

Val Glu Arg Trp Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn
545   550   555   560

Tyr Leu Phe Cys Arg Val Arg Lys Asn Gly Val Ala Ala Pro Ser Ala
                565   570   575

Thr Ser Gln Leu Ser Thr Arg Ala Leu Glu Gly Ile Phe Glu Ala Thr
            580   585   590

His Arg Leu Ile Tyr Gly Ala Lys Asp Asp Ser Gly Gln Arg Tyr Leu
        595   600   605

Ala Trp Ser Gly His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala
    610   615   620

Arg Ala Gly Val Ser Ile Pro Glu Ile Met Gln Ala Gly Gly Trp Thr
625   630   635   640

Asn Val Asn Ile Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr
                645   650   655

Gly Ala Met Val Arg Leu Leu Glu Asp Gly Asp Gly Ile Glu Gly Arg
            660   665   670

Gly Ser Ala Trp Arg His Pro Gln Phe Gly Gly
        675   680

<210> 22
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:synthetic TAT
      protein

<400> 22
Ala Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg
  1   5   10

<210> 23
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic TAT
      protein

<400> 23
Tyr Ala Arg Lys Ala Arg Arg Gln Ala Arg Arg
1 5 10

<210> 24
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic TAT protein

<400> 24
Tyr Ala Arg Ala Ala Ala Arg Gln Ala Arg Ala
1 5 10

<210> 25
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic TAT protein

<400> 25
Tyr Ala Arg Ala Ala Arg Arg Ala Ala Arg Arg
1 5 10

<210> 26
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic TAT protein

<400> 26
Tyr Ala Arg Ala Ala Arg Arg Ala Ala Arg Ala
1 5 10

<210> 27
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic TAT protein

<400> 27
Tyr Ala Arg Arg Arg Arg Arg Arg Arg Arg Arg
1 5 10

<210> 28
<211> 11
<212> PRT

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic TAT protein

<400> 28
Tyr Ala Ala Ala Arg Arg Arg Arg Arg Arg Arg
 1   5   10

<210> 29
<211> 4960
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: vector pCMV-I-Cre-pA

<400> 29
aaacagtccg atgtacgggc cagatatacg cgttgacatt gattattgac tagttattaa 60
tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg cgttacataa 120
cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt gacgtcaata 180
atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca atgggtggac 240
tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc aagtacgccc 300
cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta catgacctta 360
tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac catggtgatg 420
cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg atttccaagt 480
ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg ggactttcca 540
aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt acggtgggag 600
gtctatataa gcagagctct ctggctaact agagaaccca ctgcttactg gcttatcgaa 660
attaatacga ctcactatag ggagacccaa gctgactcta gacttaatta agcgttgggg 720
tgagtactcc ctctcaaaag cgggcatgac ttctgcgcta agattgtcag tttccaaaaa 780
cgaggaggat ttgatattca cctggcccgc ggtgatgcct ttgagggtgg ccgcgtccat 840
ctggtcagaa aagacaatct ttttgttgtc aagcttgagg tgtggcaggc ttgagatctg 900
gccatacact tgagtgacat tgacatccac tttgcctttc tctccacagg tgtccactcc 960
cagggcggcc tcgaccatgc ccaagaagaa gaggaaggtg tccaatttac tgaccgtaca 1020
ccaaaatttg cctgcattac cggtcgatgc aacgagtgat gaggttcgca agaacctgat 1080
ggacatgttc agggatcgcc aggcgttttc tgagcatacc tggaaaatgc ttctgtccgt 1140
ttgccggtcg tgggcggcat ggtgcaagtt gaataaccgg aaatggtttc ccgcagaacc 1200
tgaagatgtt cgcgattatc ttctatatct tcaggcgcgc ggtctggcag taaaaactat 1260
ccagcaacat ttgggccagc taaacatgct tcatcgtcgg tccgggctgc cacgaccaag 1320
tgacagcaat gctgtttcac tggttatgcg gcggatccga aaagaaaacg ttgatgccgg 1380
tgaacgtgca aaacaggctc tagcgttcga acgcactgat ttcgaccagg ttcgttcact 1440
catggaaaat agcgatcgct gccaggatat acgtaatctg gcatttctgg ggattgctta 1500
taacaccctg ttacgtatag ccgaaattgc caggatcagg gttaaagata tctcacgtac 1560
tgacggtggg agaatgttaa tccatattgg cagaacgaaa acgctggtta gcaccgcagg 1620
tgtagagaag gcacttagcc tgggggtaac taaactggtc gagcgatgga tttccgtctc 1680
tggtgtagct gatgatccga ataactacct gttttgccgg gtcagaaaaa atggtgttgc 1740
cgcgccatct gccaccagcc agctatcaac tcgcgccctg gaagggattt ttgaagcaac 1800
tcatcgattg atttacggcg ctaaggatga ctctggtcag agatacctgg cctggtctgg 1860
acacagtgcc cgtgtcggag ccgcgcgaga tatggcccgc gctggagttt caataccgga 1920
gatcatgcaa gctggtggct ggaccaatgt aaatattgtc atgaactata tccgtaacct 1980
ggatagtgaa acaggggcaa tggtgcgcct gctggaagat ggcgattagc cattaacgcg 2040
taaatgattg cagatccact agttctaggg ccgcgtcgac ctcgagatcc aggcgcggat 2100
caataaaaga tcattatttt caatagatct gtgtgttggt tttttgtgtg ccttgggggga 2160
gggggaggcc agaatgaggc gcggccaagg gggagggga ggccagaatg accttgggggg 2220
agggggaggc cagaatgacc ttgggggagg gggaggccag aatgaggcgc gcccccgggt 2280
accgagctcg aattcactgg ccgtcgtttt acaacgtcgt gactgggaaa accctggcgt 2340
tacccaactt aatcgccttg cagcacatcc ccctttcgcc agctggcgta atagcgaaga 2400
ggcccgcacc gatcgccctt cccaacagtt gcgcagcctg aatggcgaat ggcgcctgat 2460
gcggtatttt ctccttacgc atctgtgcgg tatttcacac cgcatatggt gcactctcag 2520
tacaatctgc tctgatgccg catagttaag ccagccccga cacccgccaa cacccgctga 2580

```
cgcgccctga cgggcttgtc tgctcccggc atccgcttac agacaagctg tgaccgtctc 2640
cgggagctgc atgtgtcaga ggttttcacc gtcatcaccg aaacgcgcga gacgaaaggg 2700
cctcgtgata cgcctatttt tataggttaa tgtcatgata ataatggttt cttagacgtc 2760
aggtggcact tttcggggaa atgtgcgcgg aacccctatt tgtttatttt tctaaataca 2820
ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa 2880
aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt ttgcggcatt 2940
ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca 3000
gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag 3060
ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc tatgtggcgc 3120
ggtattatcc cgtattgacg ccgggcaaga gcaactcggt cgccgcatac actattctca 3180
gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt 3240
aagagaatta tgcagtgctg ccataaccat gagtgataac actgcggcca acttacttct 3300
gacaacgatc ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt 3360
aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg acgagcgtga 3420
caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg gcgaactact 3480
tactctagct tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc 3540
acttctgcgc tcggcccttc cggctggctg gtttattgct gataaatctg gagccggtga 3600
gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt 3660
agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac agatcgctga 3720
gataggtgcc tcactgatta agcattggta actgtcagac caagtttact catatatact 3780
ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga tcctttttga 3840
taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgt 3900
agaaaagatc aaaggatctt cttgagatcc tttttttctg cgcgtaatct gctgcttgca 3960
aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg gatcaagagc taccaactct 4020
ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc ttctagtgta 4080
gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct 4140
aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttggactc 4200
aagacgatag ttaccggata aggcgcagcg gtcgggctga acggggggtt cgtgcacaca 4260
gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agctatgaga 4320
aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg 4380
aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggtatcttt atagtcctgt 4440
cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggcggag 4500
cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt gctggccttt 4560
tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt 4620
tgagtgagct gataccgctc gccgcagccg aacgaccgag cgcagcgagt cagtgagcga 4680
ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta 4740
atgcagctgg cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa 4800
tgtgagttag ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat 4860
gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta 4920
cgccaagcta gcccgggcta gcttgcatgc ctgcaggttt                       4960
```

<210> 30
<211> 7332
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: vector
pCMV-I-beta-pA

<400> 30

```
aaacagtccg atgtacgggc cagatatacg cgttgacatt gattattgac tagttattaa 60
tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg cgttacataa 120
cttacggtaa atggcccgcc tggctgaccg cccaacgacc cccgcccatt gacgtcaata 180
atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca atgggtggac 240
tatttacggt aaactgccca cttggcagta catcaagtgt atcatatgcc aagtacgccc 300
cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta catgacctta 360
tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac catggtgatg 420
cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg atttccaagt 480
ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg ggactttcca 540
aaatgtcgta acaactccgc cccattgacg caaatgggcg gtaggcgtgt acggtgggag 600
gtctatataa gcagagctct ctggctaact agagaaccca ctgcttactg gcttatcgaa 660
attaatacga ctcactatag ggagacccaa gctgactcta gacttaatta agcgttgggg 720
```

```
tgagtactcc ctctcaaaag cgggcatgac ttctgcgcta agattgtcag tttccaaaaa 780
cgaggaggat ttgatattca cctggcccgc ggtgatgcct ttgagggtgg ccgcgtccat 840
ctggtcagaa aagacaatct ttttgttgtc aagcttgagg tgtggcaggc ttgagatctg 900
gccatacact tgagtgacat tgacatccac tttgcctttc tctccacagg tgtccactcc 960
cagggcggcc gcaattcccg ggatcgaaa gagcctgcta aagcaaaaaa gaagtcacca 1020
tgtcgtttac tttgaccaac aagaacgtga ttttcgttgc cggtctggga ggcattggtc 1080
tggacaccag caaggagctg ctcaagcgcg atcccgtcgt tttacaacgt cgtgactggg 1140
aaaaccctgg cgttacccaa cttaatcgcc ttgcagcaca tccccctttc gccagctggc 1200
gtaatagcga agaggcccgc accgatcgcc cttcccaaca gttgcgcagc ctgaatggcg 1260
aatggcgctt tgcctggttt ccggcaccag aagcggtgcc ggaaagctgg ctggagtgcg 1320
atcttcctga ggccgatact gtcgtcgtcc cctcaaactg gcagatgcac ggttacgatg 1380
cgcccatcta caccaacgta acctatccca ttacggtcaa tccgccgttt gttcccacgg 1440
agaatccgac gggttgttac tcgctcacat ttaatgttga tgaaagctgg ctacaggaag 1500
gccagacgcg aattattttt gatggcgtta actcggcgtt tcatctgtgg tgcaacgggc 1560
gctgggtcgg ttacggccag gacagtcgtt tgccgtctga atttgacctg agcgcatttt 1620
tacgcgccgg agaaaaccgc ctcgcggtga tggtgctgcg ttggagtgac ggcagttatc 1680
tggaagatca ggatatgtgg cggatgagcg gcattttccg tgacgtctcg ttgctgcata 1740
aaccgactac acaaatcagc gatttccatg ttgccactcg ctttaatgat gatttcagcc 1800
gcgctgtact ggaggctgaa gttcagatgt gcggcgagtt gcgtgactac ctacgggtaa 1860
cagtttcttt atggcagggt gaaacgcagg tcgccagcgg caccgcgcct ttcggcggtg 1920
aaattatcga tgagcgtggt ggttatgccg atcgcgtcac actacgtctg aacgtcgaaa 1980
acccgaaact gtggagcgcc gaaatcccga atctctatcg tgcggtggtt gaactgcaca 2040
ccgccgacgg cacgctgatt gaagcagaag cctgcgatgt cggtttccgc gaggtgcgga 2100
ttgaaaatgg tctgctgctg ctgaacggca agccgttgct gattcgaggc gttaaccgtc 2160
acgagcatca tcctctgcat ggtcaggtca tggatgagca gacgatggtg caggatatcc 2220
tgctgatgaa gcagaacaac tttaacgccg tgcgctgttc gcattatccg aaccatccgc 2280
tgtggtacac gctgtgcgac cgctacggcc tgtatgtggt ggatgaagcc aatattgaaa 2340
cccacggcat ggtgccaatg aatcgtctga ccgatgatcc gcgctggcta ccggcgatga 2400
gcgaacgcgt aacgcgaatg gtgcagcgcg atcgtaatca cccgagtgtg atcatctggt 2460
cgctggggaa tgaatcaggc cacggcgcta atcacgacgc gctgtatcgc tggatcaaat 2520
ctgtcgatcc ttcccgcccg gtgcagtatg aaggcggcgg agccgacacc acggccaccg 2580
atattatttg cccgatgtac gcgcgcgtgg atgaagacca gcccttcccg gctgtgccga 2640
aatggtccat caaaaaatgg ctttcgctac ctggagagac gcgcccgctg atcctttgcg 2700
aatacgccca cgcgatgggt aacagtcttg gcggtttcgc taaatactgg caggcgtttc 2760
gtcagtatcc ccgtttacag ggcggcttcg tctgggactg ggtggatcag tcgctgatta 2820
aatatgatga aaacggcaac ccgtggtcgg cttacggcgg tgattttggc gatacgccga 2880
acgatcgcca gttctgtatg aacggtctgg tctttgccga ccgcacgccg catccagcgc 2940
tgacggaagc aaaacaccag cagcagtttt tccagttccg tttatccggg caaaccatcg 3000
aagtgaccag cgaatacctg ttccgtcata gcgataacga gctcctgcac tggatggtgg 3060
cgctggatgg taagccgctg gcaagcggtg aagtgcctct ggatgtcgct ccacaaggta 3120
aacagttgat tgaactgcct gaactaccgc agccggagag cgccgggcaa ctctggctca 3180
cagtacgcgt agtgcaaccg aacgcgaccg catggtcaga agccgggcac atcagcgcct 3240
ggcagcagtg gcgtctggcg gaaaacctca gtgtgacgct ccccgccgcg tcccacgcca 3300
tcccgcatct gaccaccagc gaaatggatt tttgcatcga gctgggtaat aagcgttggc 3360
aatttaaccg ccagtcaggc tttctttcac agatgtggat tggcgataaa aaacaactgc 3420
tgacgccgct gcgcgatcag ttcacccgtg caccgctgga taacgacatt ggcgtaagtg 3480
aagcgacccg cattgaccct aacgcctggg tcgaacgctg gaaggcggcg ggccattacc 3540
aggccgaagc agcgttgttg cagtgcacgg cagatacact tgctgatgcg gtgctgatta 3600
cgaccgctca cgcgtggcag catcagggga aaaccttatt tatcagccgg aaaacctacc 3660
ggattgatgg tagtggtcaa atggcgatta ccgttgatgt tgaagtggcg agcgatacac 3720
cgcatccggc gcggattggc ctgaactgcc agctggcgca ggtagcagag cgggtaaact 3780
ggctcggatt agggccgcaa gaaaactatc ccgaccgcct tactgccgcc tgttttgacc 3840
gctgggatct gccattgtca gacatgtata ccccgtacgt cttcccgagc gaaaacggtc 3900
tgcgctgcgg gacgcgcgaa ttgaattatg gcccacacca gtggcgcggc gacttccagt 3960
tcaacatcag ccgctacagt caacagcaac tgatggaaac cagccatcgc catctgctgc 4020
acgcggaaga aggcacatgg ctgaatatcg acggtttcca tatgggggatt ggtggcgacg 4080
actcctggag cccgtcagta tcggcggaat tacagctgag cgccggtcgc taccattacc 4140
agttggtctg gtgtcaaaaa taataataac cgggcaggcc atgtctgccc gtatttcgcg 4200
taaggaaatc cattatgtac tatttaaaaa acacaaactt ttggatgttc ggtttattct 4260
ttttctttta ctttttatc atgggagcct acttcccgtt tttcccgatt tggctacatg 4320
acatcaacca tatcagcaaa agtgatacgg gtattatttt tgccgctatt tctctgttct 4380
cgctattatt ccaaccgctg tttggtctgc tttctgacaa actcggcctc gactctaggc 4440
ggccgcgtcg acctcgagat ccaggcgcgg atcaataaaa gatcattatt ttcaatagat 4500
ctgtgtgttg gtttttgtg tgccttgggg gagggggagg ccagaatgag gcgcggccaa 4560
```

gggggagggg gaggccagaa tgaccttggg ggagggggag gccagaatga ccttggggga 4620
gggggaggcc agaatgaggc gcgcccccgg gtaccgagct cgaattcact ggccgtcgtt 4680
ttacaacgtc gtgactggga aaaccctggc gttacccaac ttaatcgcct tgcagcacat 4740
ccccctttcg ccagctggcg taatagcgaa gaggcccgca ccgatcgccc ttcccaacag 4800
ttgcgcagcc tgaatggcga atggcgcctg atgcggtatt ttctccttac gcatctgtgc 4860
ggtatttcac accgcatatg gtgcactctc agtacaatct gctctgatgc cgcatagtta 4920
agccagcccc gacacccgcc aacacccgct gacgcgccct gacgggcttg tctgctcccg 4980
gcatccgctt acagacaagc tgtgaccgtc tccgggagct gcatgtgtca gaggttttca 5040
ccgtcatcac cgaaacgcgc gagacgaaag ggcctcgtga tacgcctatt tttataggtt 5100
aatgtcatga taataatggt ttcttagacg tcaggtggca cttttcgggg aaatgtgcgc 5160
ggaaccccta tttgtttatt tttctaaata cattcaaata tgtatccgct catgagacaa 5220
taaccctgat aaatgcttca ataatattga aaaaggaaga gtatgagtat tcaacatttc 5280
cgtgtcgccc ttattccctt ttttgcggca ttttgccttc ctgtttttgc tcacccagaa 5340
acgctggtga aagtaaaaga tgctgaagat cagttgggtg cacgagtggg ttacatcgaa 5400
ctggatctca acagcggtaa gatccttgag agttttcgcc ccgaagaacg ttttccaatg 5460
atgagcactt ttaaagttct gctatgtggc gcggtattat cccgtattga cgccgggcaa 5520
gagcaactcg gtcgccgcat acactattct cagaatgact tggttgagta ctcaccagtc 5580
acagaaaagc atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc 5640
atgagtgata acactgcggc caacttactt ctgacaacga tcggaggacc gaaggagcta 5700
accgcttttt tgcacaacat gggggatcat gtaactcgcc ttgatcgttg ggaaccggag 5760
ctgaatgaag ccataccaaa cgacgagcgt gacaccacga tgcctgtagc aatggcaaca 5820
acgttgcgca aactattaac tggcgaacta cttactctag cttcccggca acaattaata 5880
gactggatgg aggcggataa agttgcagga ccacttctgc gctcggccct tccggctggc 5940
tggtttattg ctgataaatc tggagccggt gagcgtgggt ctcgcggtat cattgcagca 6000
ctggggccag atggtaagcc ctcccgtatc gtagttatct acacgacggg gagtcaggca 6060
actatggatg aacgaaatag acagatcgct gagataggtg cctcactgat taagcattgg 6120
taactgtcag accaagttta ctcatatata ctttagattg atttaaaact tcatttttaa 6180
tttaaaagga tctaggtgaa gatccttttt gataatctca tgaccaaaat cccttaacgt 6240
gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat 6300
cctttttttc tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcggtg 6360
gtttgtttgc cggatcaaga gctaccaact ctttttccga aggtaactgg cttcagcaga 6420
gcgcagatac caaatactgt ccttctagtg tagccgtagt taggccacca cttcaagaac 6480
tctgtagcac cgcctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt 6540
ggcgataagt cgtgtcttac cgggttggac tcaagacgat agttaccgga taaggcgcag 6600
cggtcgggct gaacgggggg ttcgtgcaca cagcccagct tggagcgaac gacctacacc 6660
gaactgagat acctacagcg tgagctatga gaaagcgcca cgcttcccga agggagaaag 6720
gcggacaggt atccggtaag cggcagggtc ggaacaggag agcgcacgag ggagcttcca 6780
gggggaaacg cctggtatct ttatagtcct gtcgggtttc gccacctctg acttgagcgt 6840
cgatttttgt gatgctcgtc aggggggcgg agcctatgga aaaacgccag caacgcggcc 6900
tttttacggt tcctggcctt ttgctggcct tttgctcaca tgttctttcc tgcgttatcc 6960
cctgattctg tggataaccg tattaccgcc tttgagtgag ctgataccgc tcgccgcagc 7020
cgaacgaccg agcgcagcga gtcagtgagc gaggaagcgg aagagcgccc aatacgcaaa 7080
ccgcctctcc ccgcgcgttg gccgattcat taatgcagct ggcacgacag gtttcccgac 7140
tggaaagcgg gcagtgagcg caacgcaatt aatgtgagtt agctcactca ttaggcaccc 7200
caggctttac actttatgct tccggctcgt atgttgtgtg gaattgtgag cggataacaa 7260
tttcacacag gaaacagcta tgaccatgat tacgccaagc tagcccgggc tagcttgcat 7320
gcctgcaggt tt 7332

<210> 31
<211> 72
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 31
atgccatggg ctacggccgc aagaagcgcc gccaacgccg ccgcggcatg tccaatttac 60
tgaccgtaca cc 72

<210> 32
<211> 25

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 32
tttcggatcc gccgcataac cagtg 25

<210> 33
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 33
tatatctaga ccatgggcta cggccgcaag aagc 34

<210> 34
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 34
gctaccacga ccttcgatac catcgccatc ttccagcagg cgc 43

<210> 35
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 35
taactagcgg ccgcatgtcc aatttactga ccgtacac 38

<210> 36
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 36
tcgagcggcc gccatcgcca tcttccagca ggcg 34

<210> 37
<211> 32
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 37
tatatctaga catatgacct ctcgccgctc cg 32

<210> 38
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 38
ttccgaagac gacgaaacac c 21

<210> 39
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 39
tatattcgaa gcttattaac caccgaactg cg 32

<210> 40
<211> 4847
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: vector
pGK-cre-pA

<400> 40
aggtggcact tttcggggaa atgtgcgcgg aacccctatt tgtttatttt tctaaataca 60
ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa 120
aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt ttgcggcatt 180
ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca 240
gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag 300
ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc tatgtggcgc 360
ggtattatcc cgtattgacg ccgggcaaga gcaactcggt cgccgcatac actattctca 420
gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt 480
aagagaatta tgcagtgctg ccataaccat gagtgataac actgcggcca acttacttct 540
gacaacgatc ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt 600
aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg acgagcgtga 660
caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg gcgaactact 720
tactctagct tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc 780
acttctgcgc tcggcccttc cggctggctg gtttattgct gataaatctg gagccggtga 840
gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt 900
agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac agatcgctga 960
gataggtgcc tcactgatta agcattggta actgtcagac caagtttact catatatact 1020
ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga tcctttttga 1080
taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgt 1140
agaaaagatc aaaggatctt cttgagatcc tttttttctg cgcgtaatct gctgcttgca 1200
aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg gatcaagagc taccaactct 1260
ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc ttctagtgta 1320
gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct 1380
aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttggactc 1440

aagacgatag ttaccggata aggcgcagcg gtcgggctga acggggggtt cgtgcacaca 1500
gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agctatgaga 1560
aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg 1620
aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggtatcttt atagtcctgt 1680
cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggcggag 1740
cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt gctggccttt 1800
tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt 1860
tgagtgagct gataccgctc gccgcagccg aacgaccgag cgcagcgagt cagtgagcga 1920
ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta 1980
atgcagctgg cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa 2040
tgtgagttag ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat 2100
gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta 2160
cgccaagcgc gcaattaacc ctcactaaag ggaacaaaag ctgggtaccg ggccccccct 2220
cgaggtcgac ggtatcgata agcttgatat cgaattctac cgggtagggg aggcgctttt 2280
cccaaggcag tctggagcat gcgctttagc agccccgctg gcacttggcg ctacacaagt 2340
ggcctctggc ctcgcacaca ttccacatcc accggtagcg ccaaccggct ccgttctttg 2400
gtggcccctt cgcgccactt ctactcctcc cctagtcagg aagtttcccc cagcaagctc 2460
gcgtcgtgca ggacgtgaca aatggaagta gcacgtctca ctagtctcgt gcagatggac 2520
agcaccgctg agcaatggaa gcgggtaggc ctttggggca gcggccaata gcagctttgt 2580
tccttcgctt tctgggctca gaggctggga aggggtgggt ccgggggcgg gctcaggggc 2640
gggctcaggg gcggggcggc gcccgaaggt cctcccgagg cccggcattc tgcacgcttc 2700
aaaagcgcac gtctgccgcg ctgttctcct cttcctcatc tccgggcctt tcgacctgca 2760
gctcgaggtc gaccatgccc aagaagaaga ggaaggtgtc caatttactg accgtacacc 2820
aaaatttgcc tgcattaccg gtcgatgcaa cgagtgatga ggttcgcaag aacctgatgg 2880
acatgttcag ggatcgccag gcgttttctg agcatacctg gaaaatgctt ctgtccgttt 2940
gccggtcgtg ggcggcatgg tgcaagttga ataaccggaa atggtttccc gcagaacctg 3000
aagatgttcg cgattatctt ctatatcttc aggcgcgcgg tctggcagta aaaactatcc 3060
agcaacattt gggccagcta aacatgcttc atcgtcggtc cgggctgcca cgaccaagtg 3120
acagcaatgc tgtttcactg gttatgcggc ggatccgaaa agaaaacgtt gatgccggtg 3180
aacgtgcaaa acaggctcta gcgttcgaac gcactgattt cgaccaggtt cgttcactca 3240
tggaaaatag cgatcgctgc caggatatac gtaatctggc atttctgggg attgcttata 3300
acaccctgtt acgtatagcc gaaattgcca ggatcagggt taaagatatc tcacgtactg 3360
acggtgggag aatgttaatc catattggca gaacgaaaac gctggttagc accgcaggtg 3420
tagagaaggc acttagcctg ggggtaacta aactggtcga gcgatggatt tccgtctctg 3480
gtgtagctga tgatccgaat aactacctgt tttgccgggt cagaaaaaat ggtgttgccg 3540
cgccatctgc caccagccag ctatcaactc gcgccctgga agggattttt gaagcaactc 3600
atcgattgat ttacggcgct aaggatgact ctggtcagag atacctggcc tggtctggac 3660
acagtgcccg tgtcggagcc gcgcgagata tggcccgcgc tggagtttca ataccggaga 3720
tcatgcaagc tggtggctgg accaatgtaa atattgtcat gaactatatc cgtaacctgg 3780
atagtgaaac aggggcaatg gtgcgcctgc tggaagatgg cgattagcca ttaacgcgta 3840
aatgattgca gatccactag ttctagagct cgctgatcag cctcgactgt gccttctagt 3900
tgccagccat ctgttgtttg cccctccccc gtgccttcct tgaccctgga aggtgccact 3960
cccactgtcc tttcctaata aaatgaggaa attgcatcgc attgtctgag taggtgtcat 4020
tctattctgg ggggtggggt ggggcaggac agcaaggggg aggattggga agacaatagc 4080
aggcatgctg gggatgcggt gggctctatg gcttctgagn nngaaagaac cagctggggc 4140
tcgagatcca ctagttctag cctcgaggct agagcggccg ccaccgcggt ggagctccaa 4200
ttcgccctat agtgagtcgt attacgcgcg ctcactggcc gtcgttttac aacgtcgtga 4260
ctgggaaaac cctggcgtta cccaacttaa tcgccttgca gcacatcccc ctttcgccag 4320
ctggcgtaat agcgaagagg cccgcaccga tcgcccttcc caacagttgc gcagcctgaa 4380
tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg 4440
cagcgtgacc gctacacttg ccagcgccct agcgcccgct cctttcgctt tcttcccttc 4500
ctttctcgcc acgttcgccg gctttccccg tcaagctcta aatcgggggc tccctttagg 4560
gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg gtgatggttc 4620
acgtagtggg ccatcgccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt 4680
ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggtctattc 4740
ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta 4800
acaaaaattt aacgcgaatt ttaacaaaat attaacgctt acaattt 4847

<210> 41
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 41
catctccggg cctttcgacc tg 22

<210> 42
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 42
gcgatcggtg cgggcctctt c 21

Claims

1. Use of a fusion protein comprising
(a) a site-specific DNA recombinase domain and
(b) a protein transduction domain (PTD)
for preparing an agent for inducing target gene alterations in a living organism or cell culture, wherein said living organism carries at least one or more recognition sites for said site-specific DNA recombinase integrated in an endogenous gene.

2. The use of claim 1, wherein the PTD is not derived from Antennapedia and preferably is a PTD derived from the VP22 protein of HSV or from the TAT protein of HIV.

3. Use of a fusion protein comprising
(a) a site-specific DNA recombinase domain and
(b) a protein transduction domain (PTD) being not derived from Antennapedia and preferably being derived from the VP22 protein of HSV or from the TAT protein of HIV
for preparing an agent for inducing target gene alterations in a living organism or cell culture, wherein said living organism carries at least one or more recognition sites for said site-specific DNA recombinase integrated in its genome.

4. The use of claim 3, wherein the recognition sites for said site specific recombinase is present within an endogenous gene or a transgene.

5. The use of any one of claims 2 to 4, wherein the TAT protein comprises
(i) the amino acid sequence YGRKKRRQRRR (SEQ ID NO: 10) or a mutant thereof including
(ii) peptides having the amino sequences

AGRKKRRQRRR (SEQ ID NO:22)
YARKARRQARR (SEQ ID NO:23)
YARAAARQARA (SEQ ID NO:24)
YARAARRAARR (SEQ ID NO:25)
YARAARRAARA (SEQ ID NO:26)
YARRRRRRRRR (SEQ ID NO:27)
YAAARRRRRRR (SEQ ID NO:28);
preferably the TAT protein consists of one of the sequences shown in (i) or (ii) above.

6. The use of any one of claims 2 to 4, wherein the VP22 protein comprises the amino acid 16-157 of SEQ ID NO:14.

7. The use of any one of claims 1 to 6, wherein the site-specific DNA recombinase domain is selected from a recombinase protein derived from Cre, Flp, ϕC31 recombinase, and R recombinase and preferably is Cre having amino acids 15 to 357 of SEQ ID NO: 2 or Flpe having amino acids 15 to 437 of SEQ ID NO: 4.

8. The use of any one of claims 1 to 7, wherein the protein transduction domain is fused to the N-terminal of the site-specific DNA recombinase domain.

9. The use of any one of claims 1 to 8, wherein the protein transduction domain is fused to the site-specific DNA recombinase domain through a direct chemical bond or through a linker molecule.

10. The use of any one of claim 9, wherein the linker molecule is a short peptide having 1 to 20, preferably 1 to 10 amino acid residues.

11. The use of any one of claims 1 to 10, wherein said fusion protein further comprises additional functional sequences.

12. The use of claim 1, wherein the fusion protein has the sequence shown in SEQ ID NOs: 2, 4, 6 or 8.

13. The use of any one of claims 1 to 12, wherein the living organism is a vertebrate, preferably a rodent or a fish.

14. A method for inducing gene alterations in a living organism which comprises administering to said living organism, a fusion protein comprising a site-specific DNA recombinase domain and a protein transduction domain as defined in claims 1 to 12, wherein said living organism carries at least one or more recognition sites for said site-specific DNA recombinase integrated in its genome.

15. A fusion protein comprising

(a) a site-specific DNA recombinase domain as defined in claims 2 to 9 and

(b) a protein transduction domain (PTD) as defined in claims 2 to 9

provided that when (a) is the wild-type Flp or Cre then (b) is not the full length VP22 protein of HSV.

16. The fusion of claim 15, wherein the (PTD) is derived from the TAT protein of HIV.

17. A DNA sequence coding for the fusion protein of claim 15 or 16, said DNA sequence preferably comprising the sequence shown in SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 18 and/or 20.

18. A vector comprising the DNA sequence of claim 17.

19. A host cell transformed with the vector of claim 18 and/or comprising the DNA of claim 17.

20. A method for producing the fusion protein of claim 15 which comprises culturing the transformed host cell of claim 19 and isolating the fusion protein.

21. An injectable composition comprising the fusion protein as defined in claims 1 to 12 or 15 to 16.

Fig. 1

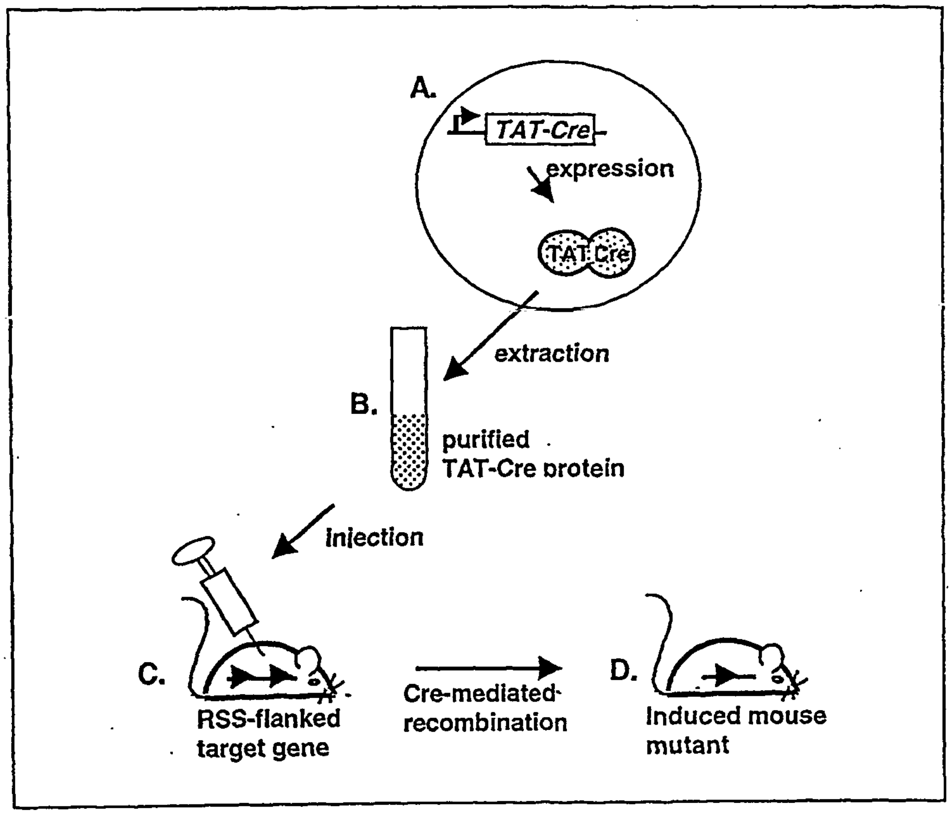
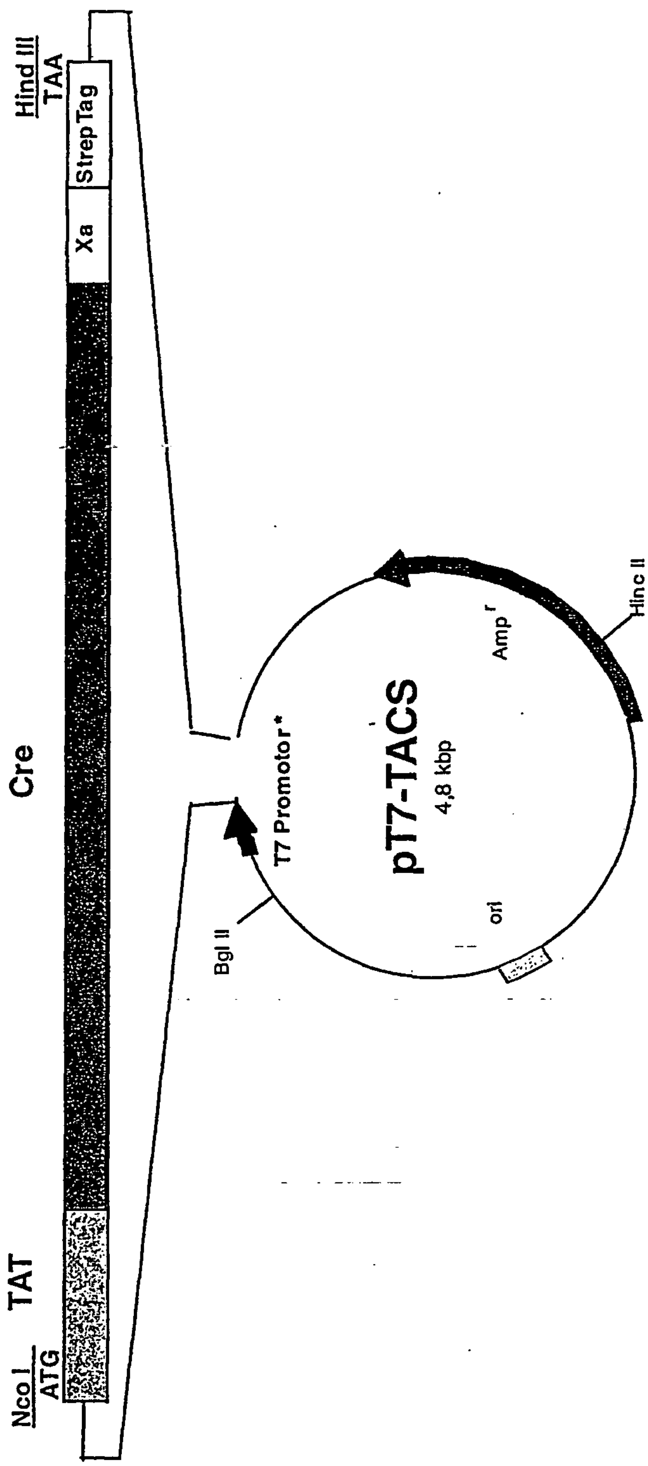
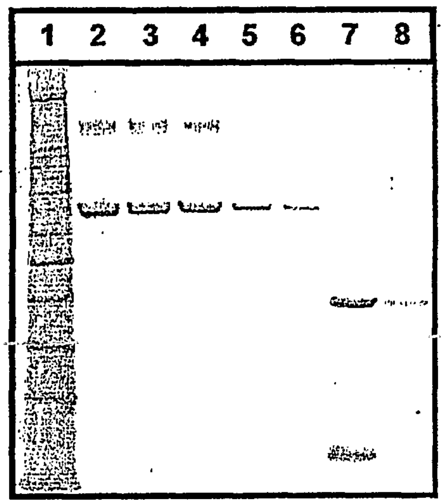

A.
TAT-Cre
expression
TAT Cre
extraction
B.
purified
TAT-Cre protein
injection
C.
RSS-flanked
target gene
Cre-mediated
recombination
D.
Induced mouse
mutant

Fig. 2


Nco I
ATG
TAT
Cre
Xa
StrepTag
Hind III
TAA
T7 Promotor*
Bgl II
pT7-TACS
4,8 kbp
Amp r
Hinc II
ori



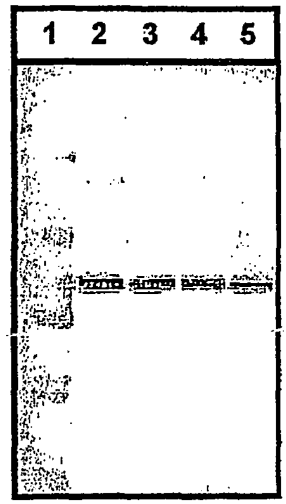


Figure 3

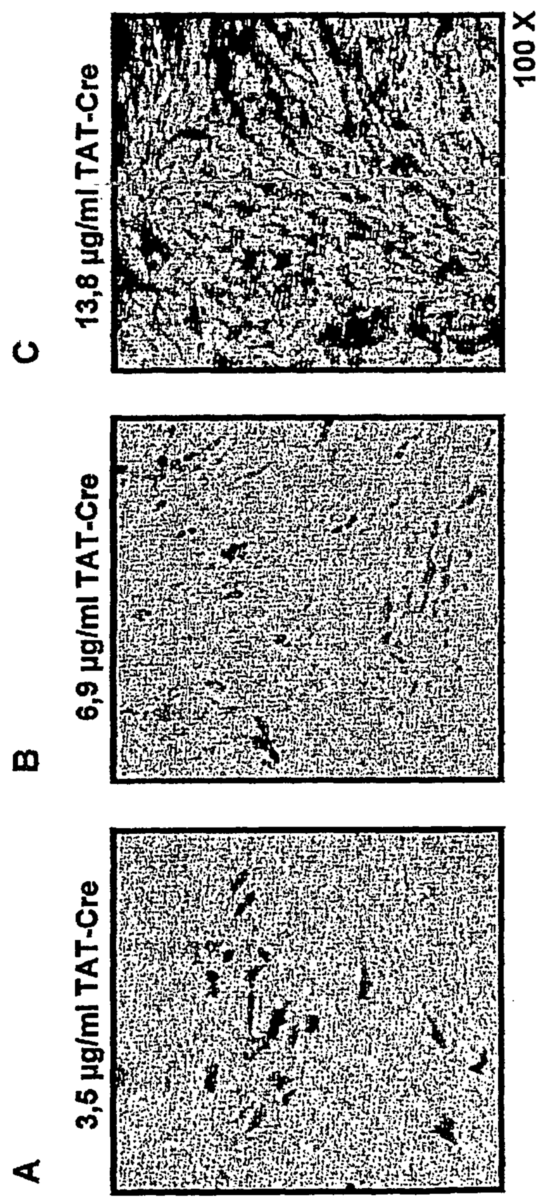
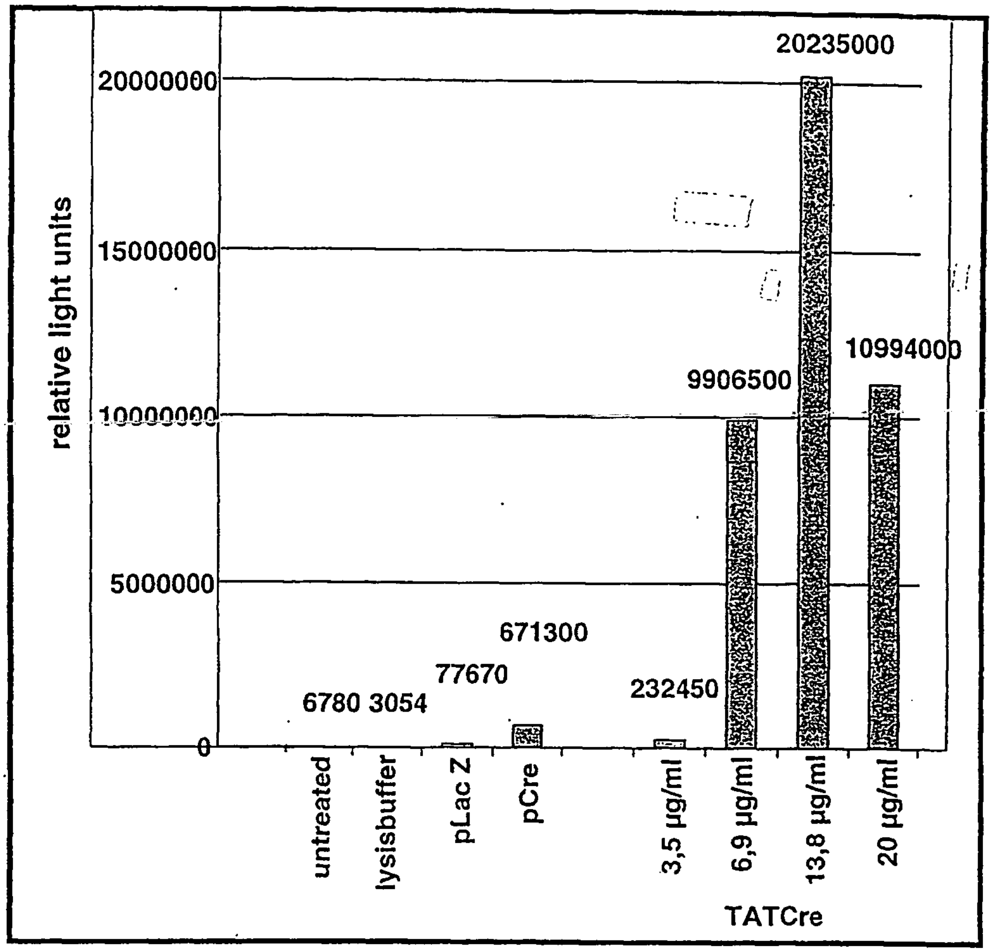


Figure 4

Fig. 5


relative light units
20000000
15000000
10000000
5000000
0
20235000
10994000
9906500
671300
77670
6780 3054
232450
untreated
lysisbuffer
pLac Z
pCre
3,5 µg/ml
6,9 µg/ml
13,8 µg/ml
20 µg/ml
TATCre

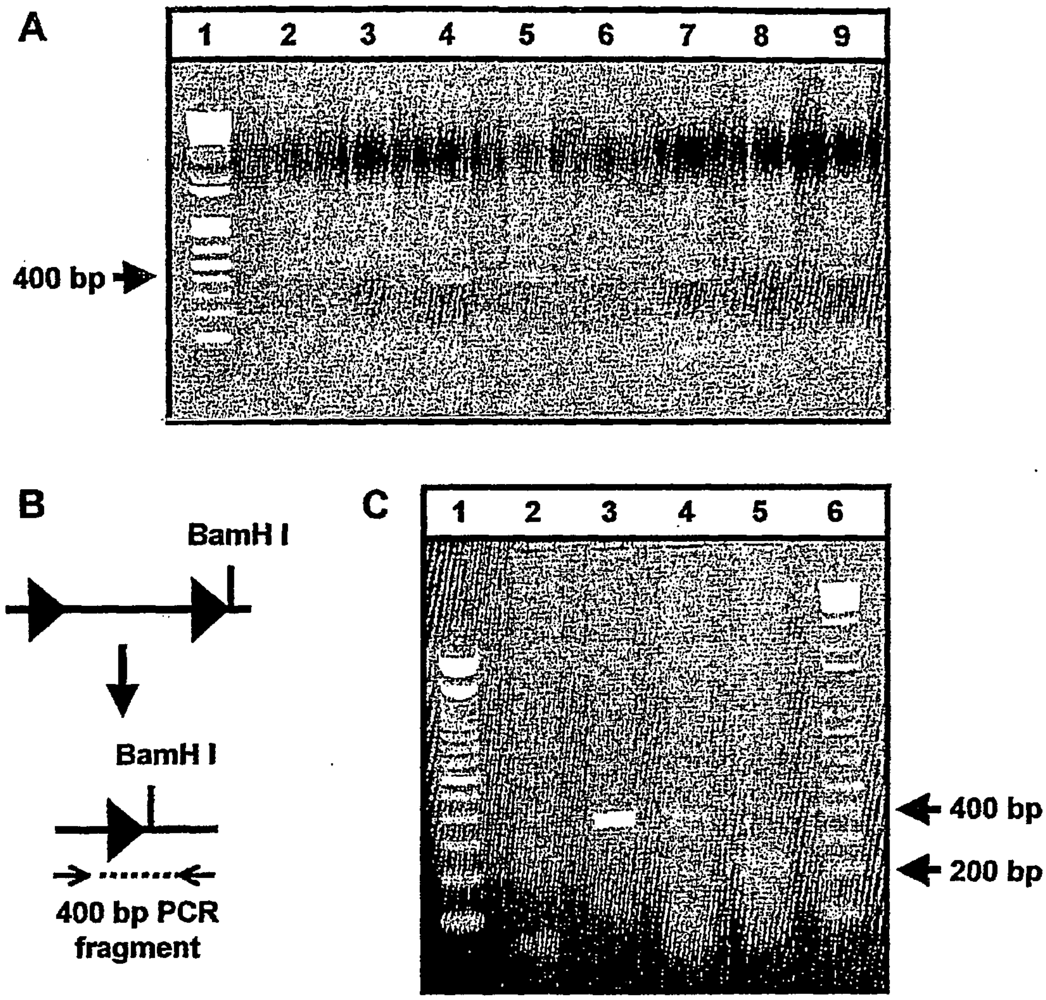
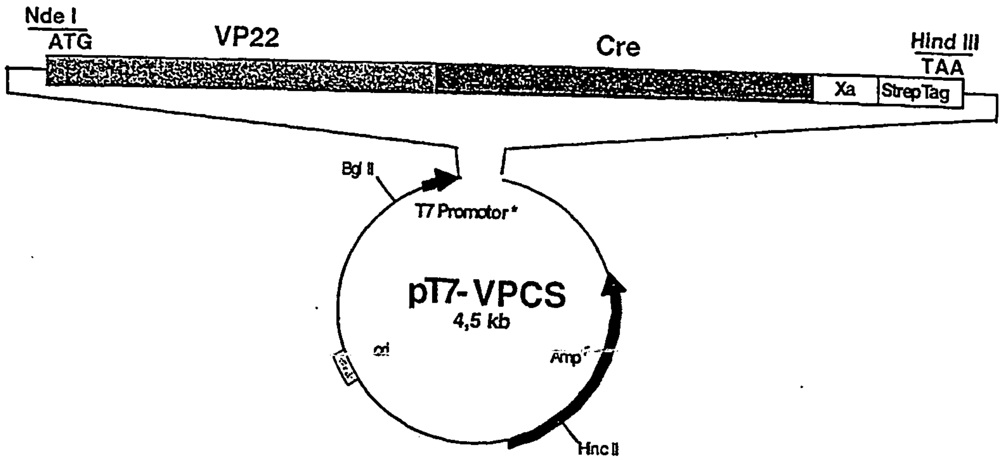

A
1 2 3 4 5 6 7 8 9
400 bp
B
BamH I
BamH I
400 bp PCR fragment
C
1 2 3 4 5 6
400 bp
200 bp

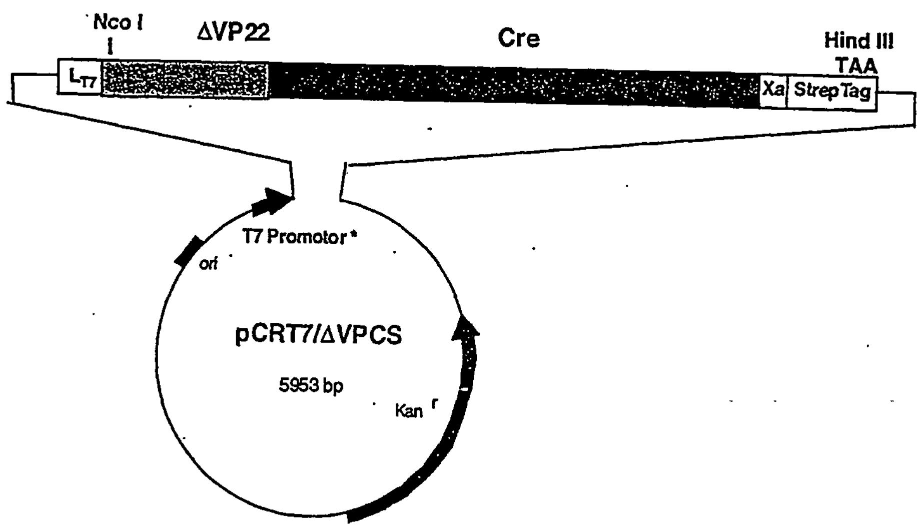
Figure 6

Fig. 7

A


Nde I
ATG
VP22
Cre
Hind III
TAA
Xa
StrepTag
Bgl II
T7 Promotor*
pT7-VPCS
4,5 kb
ori
Hnc II

B


Nco I
ΔVP22
Cre
Hind III
TAA
Xa
StrepTag
T7 Promotor*
ori
pCRT7/ΔVPCS
5953 bp
Kan r

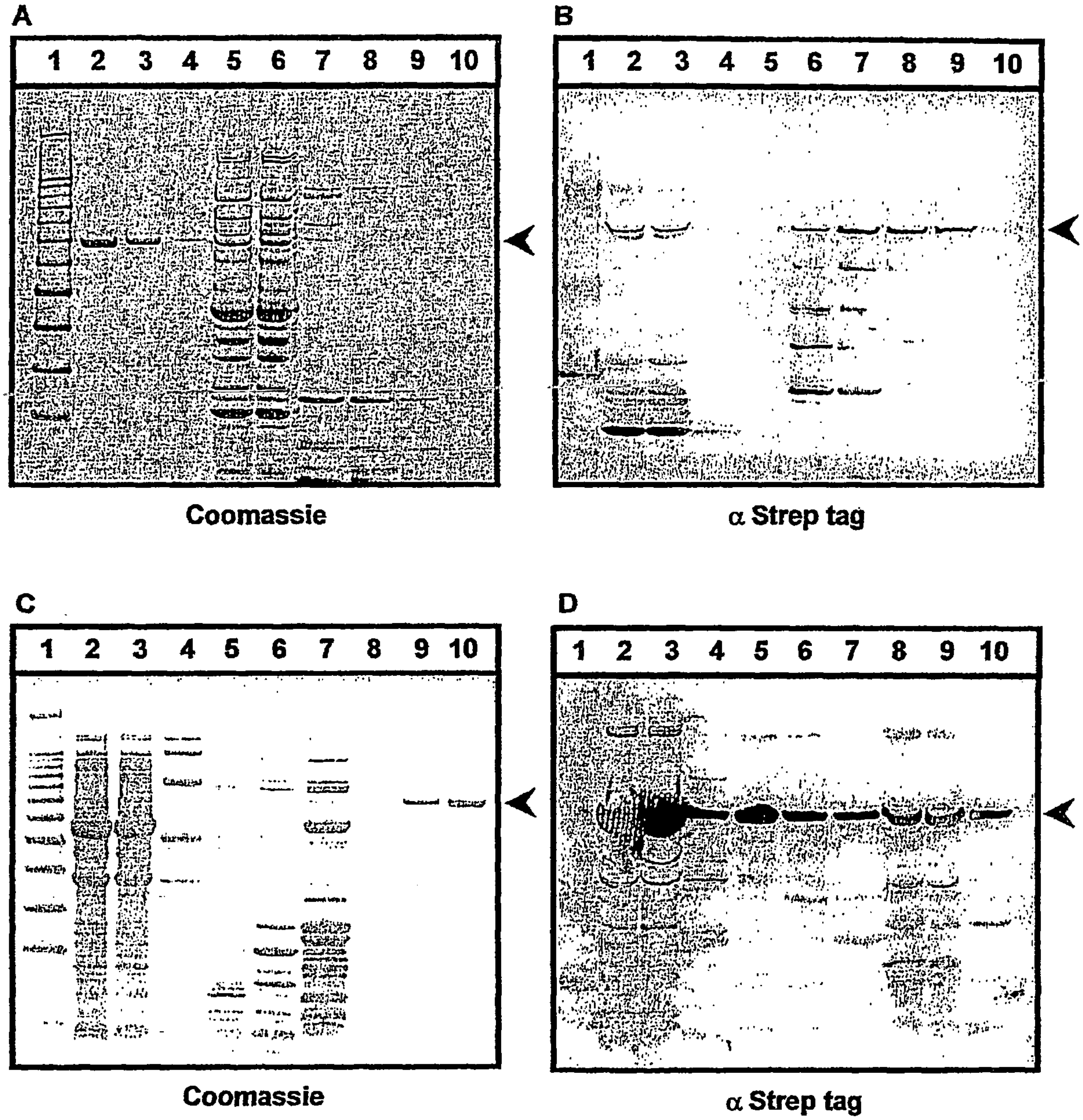

A
1 2 3 4 5 6 7 8 9 10
Coomassie
B
1 2 3 4 5 6 7 8 9 10
α Strep tag
C
1 2 3 4 5 6 7 8 9 10
Coomassie
D
1 2 3 4 5 6 7 8 9 10
α Strep tag

Figure 8

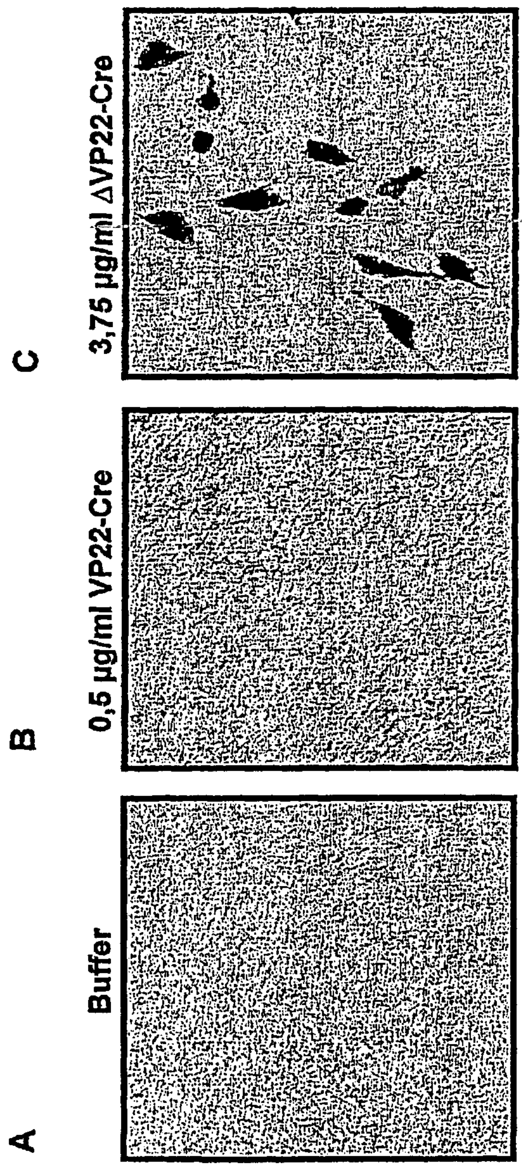
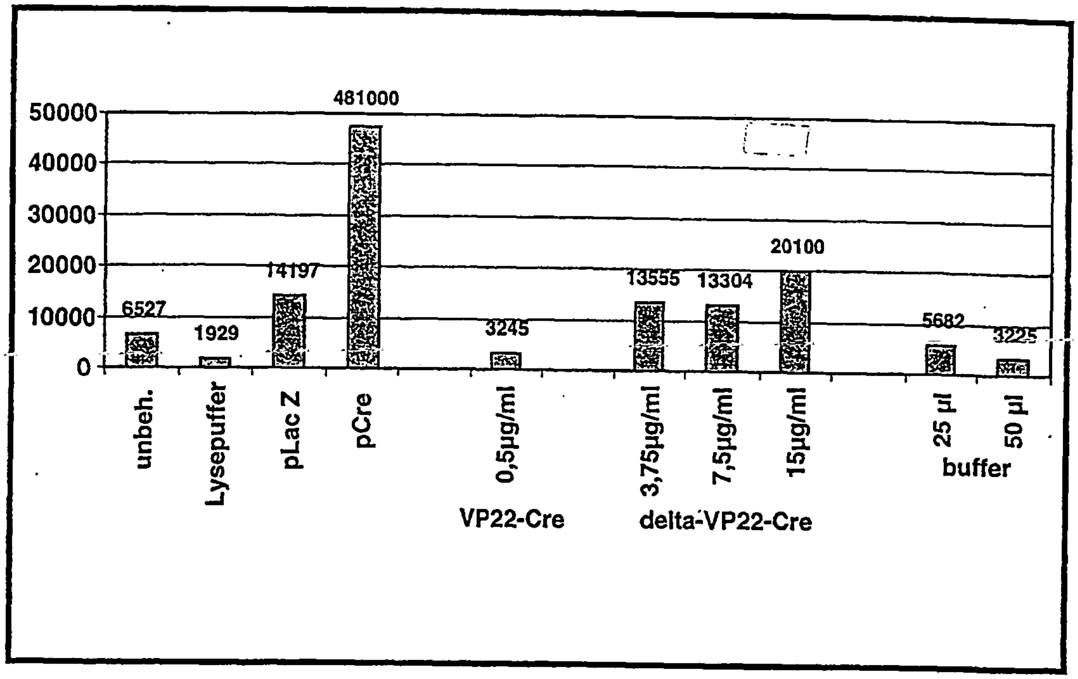


Figure 9

Fig. 10


50000
40000
30000
20000
10000
0
481000
14197
6527
1929
3245
13555
13304
20100
5682
3225
unbeh.
Lysepuffer
pLac Z
pCre
0,5µg/ml
3,75µg/ml
7,5µg/ml
15µg/ml
25 µl
50 µl
buffer
VP22-Cre
delta-VP22-Cre

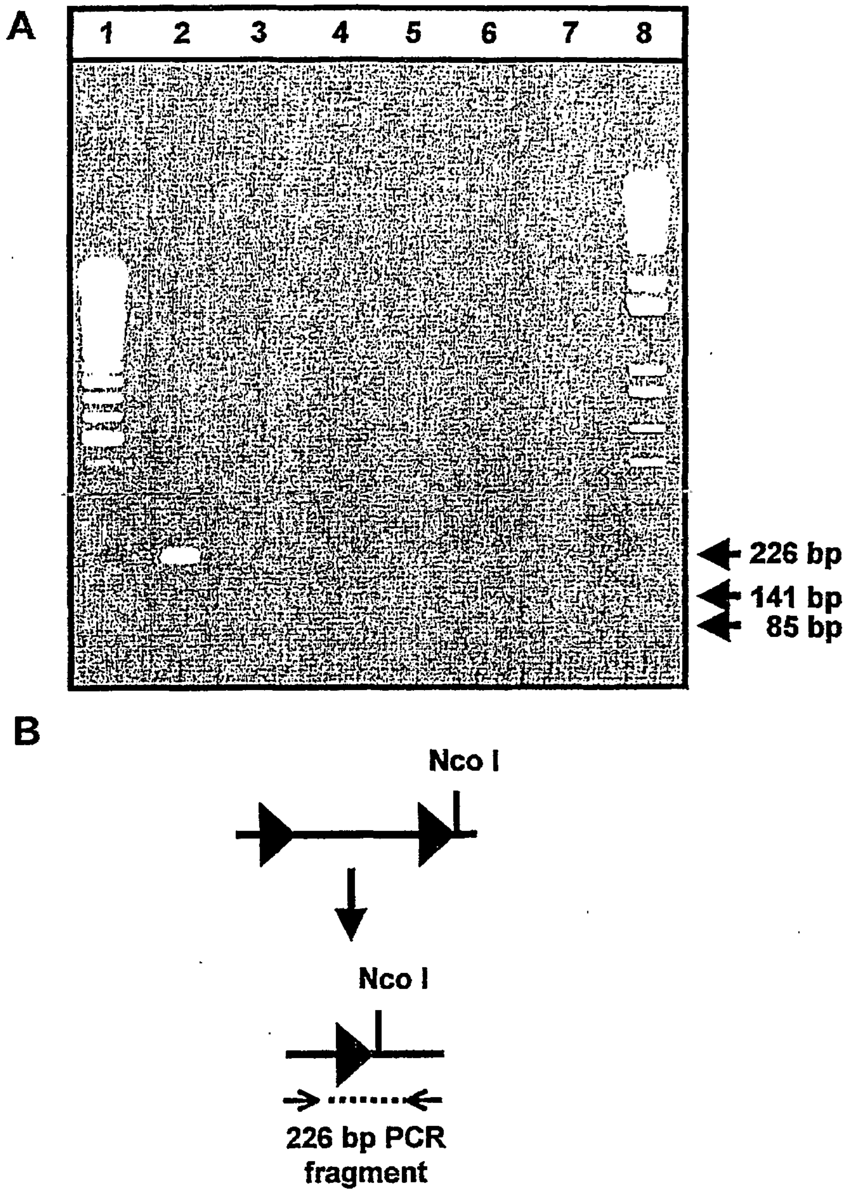


Figure 11